

EXPLORING THE ROLE OF ATTITUDES IN
NEW DIALECT FORMATION IN
HOHHOT, CHINA

by

Xuan Wang

A thesis submitted to the University of Canterbury

in fulfilment of the requirements of the degree of

Doctor of Philosophy in Linguistics

September 2017

Contents

Contents	i
List of Tables	v
List of Figures.....	vii
Acknowledgements.....	x
Abstract.....	xiv
1 Introduction.....	1
1.1 Attitude-language correlation.....	1
1.2 Attitudes in koinéization	4
1.3 Attitudes and awareness.....	6
1.4 Thesis structure	7
2 Background.....	9
2.1 Attitude-language correlation.....	9
2.1.1 Attitude and its complicated nature	10
2.1.2 Attitudes and social interaction	14
2.2 Attitudes in koinéization	16
2.3 Attitude, awareness and salience.....	23
2.4 Relevant studies in the Chinese contexts	28
3 The locality, Hohhot.....	40
3.1 Hohhot and the emergence of Hū Pǔ.....	40
3.2 Sociodemographic background	44

3.2.1 Geographic origins.....	44
3.2.2 Socio-economic status.....	45
3.3.3 Social networks, the work unit social structure	45
3.3.4 Isolation from original hometowns.....	47
3.3 Social conflicts between the New Town and the Old Town	48
3.3.1 Evidence from anthropological literature	48
3.3.2 Evidence from local online forums.....	49
4 Research design and methodology	52
4.1 Entering the community.....	52
4.2 Participants and data collection procedure	54
4.3 My own stance in the community.....	56
4.4 Linguistic variables.....	57
4.4.1 L-words.....	57
4.4.1.1 The stress pattern variable.....	59
4.4.1.2 The fricative variable	60
4.4.2 Linguistic data collection: word elicitation task.....	60
4.5 Attitudinal index scores	64
4.5.1 The questionnaire: AAS	64
4.5.2 Calculating attitudinal index scores: Principal Component Analysis	68
4.5.2.1 PCA of New Town AAS.....	69
4.5.2.2 PCA of Old Town AAS	74
4.6 Social interaction scores	77
4.6.1 Collecting social networks data	77

4.6.2 Calculating the social interaction index score	79
4.7 Demographic variables.....	80
4.8 Quantitative analysis techniques	81
5 Stress pattern variation	84
5.1 Description of the variable.....	84
5.2. Methodology.....	85
5.3. Results of the full data set.....	88
5.4. Results of the New Town data set	95
5.4.1 The effects of attitudes.....	96
5.4.2 Interaction between attitudes and social contact.....	106
5.5 Results of the Old Town data set	110
5.6 Validity of the attitudinal index scores	118
6 Fricative variation	124
6.1 Description of the variable.....	125
6.2 Methodology.....	129
6.3 Results of the full data set.....	134
6.4 Results of the New Town data set	140
6.5 Results of the Old Town data set	146
7 Discussion.....	153
7.1 Attitude as a predictor of linguistic behavior	153
7.2 Interaction between attitudes and social contact	158
7.3 The role of attitudes in koinéization.....	160

7.4 Effect of attitudes in relation to awareness in speech production	165
7.5 The stress pattern variable and the fricative variable	170
8 Conclusion	172
8.1 Limitations and future research.....	174
References	177
Appendix A: List of L-words	195
Appendix B: Slides used in the word elicitation task	197
Appendix C: The AAS questionnaires.....	205
Appendix D: Interview topics and questions	213
Appendix E: Social networks activity: the wedding guest list task.....	216
Appendix F: Informants' social information	219
Appendix G: Raw data	221
Appendix H: Human Ethics Application.....	223
Appendix I: Permission from publishers.....	235

List of Tables

Table 4.1 Participants' distribution.....	54
Table 4.2 Some examples of l-words in Jìn dialect and their corresponding monosyllabic forms	58
Table 4.3 Four factors revealed by PCA of the New Town AAS and the questions that have higher loadings on each factor	71
Table 4.4 Four principal components revealed by PCA of New Town AAS.	72
Table 4.5 Four factor scores revealed by PCA of New Town AAS.....	73
Table 4.6 Questions that have high loadings on the one factor extracted from the Old Town AAS responses.....	76
Table 4.7 Factor scores revealed by PCA of Old Town AAS.	76
Table 4.8 The calculation chart for speakers' social contact with Jìn speakers.....	80
Table 5.1 Examples of l-words that are pronounced in different stress patterns in Jìn dialect and Pǔtōnghuà.....	85
Table 5.2 Output of the best model for the stress pattern variable in the full data set.....	89
Table 5.3 Language use and language choice of the Old Town middle-aged participants.....	93
Table 5.4 Output of the best model for stress pattern in New Town data set.....	96
Table 5.5 Number of Jìn, non-Jìn and bidialectal speaker in each age group.....	111
Table 5.6 Output of the best model in the Old Town data set.....	112

Table 5.7 Attitudinal scores ATTJIN, HOHORE and ATTOT for New Town middle-aged speakers	118
Table 6.1 Four acoustic measurement factors revealed by PCA.....	133
Table 6.2 Outcome of the best model for correlation between auditory coding and acoustic measures.....	133
Table 6.3 Output of the best model for fricative variation in the full data set.	136
Table 6.4 Output of the best model for the fricative variation in the New Town data set. ..	141
Table 6.5 Output of the best model for the fricative variation in the Old Town data set.	147

List of Figures

Figure 3.1 Map of the Hohhot urban area, with the New Town and the Old Town separated by the thick dashed line.	41
Figure 3.2 Location of Inner Mongolia, and the neighboring provinces of Shānxī, Shǎnxī, and Héběi.	42
Figure 4.1 Original hometowns of the state-sponsored participants in this project, with migration movement directions indicated by the dotted arrows.	55
Figure 4.2 An example of the word elicitation task with the target l-word 划拉/xuə ⁴³ la ³⁵ / “to scribble”. Note that these were animations in the actual elicitation task.	61
Figure 4.3 Slide 5 of the word elicitation task with the target l-word 提溜/t ^h iə ⁴³ liu ³¹ / ‘carry something randomly’.	62
Figure 4.4 An example of the AAS questionnaire.	67
Figure 4.5 Scree plot from PCA of the New Town AAS.	70
Figure 4.6 Scree plot from PCA of Old Town AAS.	75
Figure 4.7 List of questions used in the interview to collect speakers’ social contact information.	78
Figure 5.1 Correlation between the vowel duration ratio and the log odds of the probability of a token being coded as weak-strong (w-s).	87
Figure 5.2 Effect of educational background (EDU) on stress pattern variation in the full data set. Y refers to speakers who have received a college education.	90
Figure 5.3 Plots of the three-way interaction between TOWN, SEX and AGE for the stress pattern variable in the full data set.	91

Figure 5.4 Plots of the two-way interactions between AGE and three attitudinal index scores in the New Town data set.	97
Figure 5.5 Effects of interaction between SOCNET and AGE in the New Town data set.....	106
Figure 5.6 Effects of the interactions between SOCNET and attitudinal scores in the New Town older speakers.....	107
Figure 5.7 Effects of interaction between SOCNET and attitudinal score HOHORE in the New Town middle-aged group.....	109
Figure 5.8 Effects of interactions between SOCNET and attitudinal score in the New Town younger generation.	110
Figure 5.9 Effects of the three-way interaction between ATT, SEX and MOVNT in the Old Town data set.....	113
Figure 5.10 Plot of the interaction between SOCNET and MOVNT in the Old Town data set.	117
Figure 6.1 Example of the labeling of plosive [p ^h] tokens.	129
Figure 6.2 Example of the labeling of fricative [h] tokens.	129
Figure 6.3 Spectrograms of [h] and [x] produced by the same speaker NMF1.	131
Figure 6.4 Acoustic spectra of the same two tokens [x] (left) and [h] (right) produced by speaker NMF1.	132
Figure 6.5 Plots of the effects of FVOWEL and the interaction between PHONEME and STRESS for the fricative variable in the full data set.....	137
Figure 6.6 Plots of the 3-way interaction between TOWN, SEX and AGE for the fricative variable in the full data set.....	138

Figure 6.7 Plots of the effects of FVOWEL and PHONEME in the New Town data set.....	142
Figure 6.8 Plot of the effect of SOCNET in the New Town data set.	142
Figure 6.9 Plots of the interaction between AGE and three attitudinal index scores in the New Town data set.....	143
Figure 6.10 Plots of the effects of FVOWEL and PHONEME in the Old Town data set.	147
Figure 6.11 Plots of the three-way interaction between ATT, SEX, and MOVNT in the Old Town data set.....	148
Figure 6.12 Plot of the interaction between AGE and SOCNET in the Old Town data set.	150

Acknowledgements

I would like to express my immense gratitude to my supervisors Kevin Watson and Lynn Clark, who guided me through the Ph.D. process. It is through their direction and supervision that I have grown from someone who had little knowledge about sociolinguistics to an independent researcher today. Their invaluable help, constructive suggestions, effective direction and invariable trust in me allowed me to develop my interests into this dissertation, and I hope into my future career. Kevin introduced me to the field of variationist sociolinguistics and led me to this thesis topic through numerous insightful discussions. He's never short of great ideas – I can't remember how many times I knocked at his door with a worried and confused look on my face and came out with full confidence and motivation to continue my research. Lynn taught me so much in every aspect, and treated me as a friend. She was my “Mrs. Motivator” in encouraging me to achieve much more than I could ever imagine. I cannot thank them enough for everything they have done for me in the past several years.

I would also like to thank all the staff and fellow students at the Linguistics Department and the New Zealand Institute of Language Brain and Behavior at the University of Canterbury, for their constant support and for being such a wonderful group. I've received considerable help at various socio meetings and feel privileged to have spent my Ph.D. years in this stimulating and fun environment. My special thanks go to Jen Hay, whose constructive comments at my Ph.D. confirmation and socio talks at different stages gave me crucial help in improving this thesis. I'd also like to thank Keyi Sun, Clay Beckner, and Darcy Rose who generously offered their time and help for me to dig into more advanced statistics in R. Vica Papp and Ryan Podlubny provided inspiring suggestions regarding the acoustic analysis of the fricative variable and amazing help with Praat. Heidi Quinn, Beth Hume, and Susan Foster-Cohen always gave me the warmest encouragement at every stage. Ksenia Gnevshcheva, Darcy Rose, Wakayo Mattingley, and Jacq Jones are the best office mates and

friends, who helped me through the dark times and made me feel like home at the office. Without their encouragement and support, I could never make it this far. Fantastic times were also had with my department mates Matthias Heynes, Andy Gibson, Daniel Bürkle, Ahmad Heider, Jiao Dan, Annalise Fletcher, Khalid Alijawazneh, Mineko Shirakawa, Keyi Sun, Arshad Ali, Ryan Podlubny, Jacqui Nokes, Moonsoon Choi, Mohammed Dagamseh, Daiki Hashimoto, and Liam Walsh. Thank you all for making my life at UC a wonderful memory. I'm also grateful to have the opportunity to meet and talk to many illustrious visitors to the institute: Katie Drager, Abby Walker, Erez Levon, Harald Baayen, and Graeme Trousdale offered me valuable suggestions to improve the thesis. Additional credit is owed to Stephanie Stokes, who initially encouraged me to come to UC and put me into contact with Kevin.

I feel incredibly lucky to have received enormous help from many Chinese professors at the School of Language, Social and Political Sciences at UC. Licheng Li listened to my early ideas and shared with me his wisdom in Chinese phonology. Hong Xiao and Zhifang Song provided valuable suggestions to my questionnaire design and fieldwork plans, and helped me to adapt them more to the Chinese contexts.

I am indebted to my friend Sarala Puthuval, who is also doing research on Hohhot and introduced me numerous books on the anthropological and historical background of Hohhot written by western scholars. Sarala's insights and help allowed me to look at Hohhot from an outsider's point of view. While collaborating on a book chapter with Sarala, I learned so much from her about academic writing and using various software tools for research, and my thesis is much better for her guidance. My thanks also go to Hongxia He and the librarians at the Inner Mongolia Library who helped me access books on the migration history of Hohhot, as well as relevant census data, both of which were integral to my thesis.

I would also like to thank Gregory Guy, James Stanford, Jeff Siegel, and Mariam Meyerhoff for their insightful suggestions to my thesis at various occasions. My thanks also extend to

Lauren Hall-Lew who agreed to be my mentor and helped me a great deal in advertising my research on social media.

I would also like to express my sincere gratitude and thanks to my family. My parents, Youlin Wang and Xinhua Ma, have been extremely supportive and proud of my academic endeavors and my research on Hohhot dialects. They shared with me their knowledge of the history of Hohhot and their valuable childhood experience concerning the local dialects, which helped me clarify and improve my early ideas about the formation of Hū Pǔ. During my fieldwork in Hohhot, they also helped immensely with participant recruitment. I also owe a big debt to my aunt, Fengqin Wang, who gave me tremendous help in recruiting older participants. During the time I stayed at her place, she cooked for me and arranged all the interviews for me in advance so that I could fully concentrate. It is with her tireless aid that I was able to complete my data collection effectively in two months. The success of my fieldwork would not have been possible without the help of all my family and friends from Hohhot. Also, I must say thank you to my husband, Weixing Sun, who moved all the way to New Zealand to support and take care of me. During the last three years, he took on the roles of chef, chauffeur and cleaner, working tirelessly when I was busy meeting the deadlines, and was always the first sounding board of my new ideas. During my research design, he volunteered to be the “actor” in my word elicitation task videos, and was so patient with all my weird requirements. As a computer engineer, he also helped me a great deal with stats and all of the technical issues I had encountered during the journey of my Ph.D. I could not have accomplished this thesis without his support and love.

I acknowledge the China Scholarship Council and UC for funding this research. The School of Language, Social, and Political Sciences also awarded me several small research grants for data collection and conference travels, which I highly appreciate.

My special thanks go to all 67 participants from my hometown Hohhot for their willingness to share their experience and interesting stories with me and opening up with whatever questions I asked. Every time I listened to their recordings, I could not help laughing

because of their good humor, which made my coding and transcribing process not dull at all. Many of them showed great interest in my research, and some have still been sending me interesting online posts they read about Hohhot dialects. Thank you all for your engagement and enthusiasm, which made this thesis possible.

Abstract

This dissertation is a variationist sociolinguistic study exploring the role of speakers' attitudes in speech production and the language change involved in large-scale dialect contact situations where a new dialect mixture is formed, by presenting the case of Hohhot (呼和浩特), a Chinese immigrant city. The study employs quantitative and qualitative methods to investigate the effects of speakers' attitudes on their linguistic behavior and the role attitudes may play in different stages of the new dialect formation process.

In Hohhot, contact between the locally-born residents who speak the local Jìn dialect (Jìn Yǔ 晋语), and migrants from all parts of China, who arrived in the 1950s and 1960s, has led to the formation of a mixed, new vernacular, known locally as “Hū Pǔ (呼普)”, which means “Hohhot Pǔtōnghuà (普通话 ‘standard Mandarin’)”. Anthropological literature of Hohhot has reported clear social stratification and intense social conflicts between the locally-born and the migrant communities (Jankowiak, 1993; Borchigud, 1996). Against this background, this thesis examines whether the linguistic variation in Hohhot people's speech is conditioned by their social attitudes, and more generally, how the formation of Hū Pǔ is influenced by such socio-psychological factors as speakers' attitudes and identities.

Data was collected in the fall of 2014 and 67 speakers from three generations were interviewed – 35 from the migrant community and 32 who were locally born. Individual speakers' attitudes and identity information were collected using overt questionnaires by employing the “Attitude Analog Scale” (Llamas & Watt, 2014), and principal component analysis was conducted to reveal underlying attitudinal categories from the responses and build up attitudinal index scores for each speaker. Individuals' social contact with Jìn speakers was also collected in the interviews, as well as other traditional demographic information such as age, sex, and education. Language production data were collected from interviews and an elicitation task designed to explore variations in a set of disyllabic lexemes known as “l-words” (Hou, 1999). Two linguistic variables were examined in l-words.

A stress pattern variable displays variation in that a weak-strong pattern is typical of Jìn dialect, and a strong-weak pattern is typical of Pǔtōnghuà. A fricative variable indicates whether the initial sounds [p^h, t^h, k^h, h] in l-words contain a period of following frication, often [x].

Mixed-effects modeling in R (R core Team, 2015) found that speakers' attitudinal index scores were significant predictors for both variables, and the effects of attitudes were still found even when speakers' social interaction with Jìn-communities was considered. Investigation of the migrant community found that the attitudinal scores representing different aspects of speakers' socio-psychological orientations had different effects in the three generations, which I argue is related to the likely change of social meanings attached to the linguistic variables – the later generations tended to adopt these Jìn features as markers of their urban Hohhot identity. The two linguistic variables examined also demonstrate different levels of awareness among speakers. Therefore, this thesis also explores the effects of attitude on speech production in relation to awareness or salience. The findings suggest that speakers' explicit awareness was not a threshold for the attitude-language correlation – speakers' attitudes were still found to be significant predictors of speakers' production even when a variable was below their conscious awareness.

Overall, the results provide evidence that speakers' overtly offered attitudes are likely to predict their linguistic behavior, and the effects of attitudes can be independent from other closely related social factors such as speakers' social networks. The findings also suggest that the role attitudes and identities play in new dialect formation can be very complex, and it is essential to understand speakers' socio-psychological orientations in specific social contexts.

1 Introduction

The work presented in this thesis is a variationist sociolinguistic study of a Chinese immigrant city, Hohhot. The main focus is to explore the role of attitudes in dialect contact situations where a new dialect mixture is formed. Specifically, it employs quantitative and qualitative methods to examine how speakers' social attitudes might correlate with, and perhaps influence, their linguistic behavior, and also how the effects of attitudes might function in different stages of the new dialect formation process.

1.1 Attitude-language correlation

In sociolinguistic studies, the role of socio-psychological factors such as speakers' attitudes and their identities is often invoked to explain why a linguistic feature has spread from one speech community to another (e.g., Stausland Johnsen, 2015), or the extent to which a speaker adopts linguistic features from his/her interlocutors (e.g., Pardo, Gibbons, Suppes, & Krauss, 2012; Stuart-Smith, Pryce, Timmins, & Gunter, 2013). Communication accommodation theory (hereafter, CAT; Giles & Coupland, 1991; Shepard, Giles, & Le Poire, 2001) advocates that speakers will make adjustments to their speech as a behavioral signal of their own attitudes. Speakers will linguistically converge to the interlocutor towards whom they have positive attitudes, while divergence will occur if they hold negative attitudes. A counter view to CAT proposed by Trudgill (2004, 2008) leaned towards a more deterministic view of accommodation in certain new dialect formation situations, claiming that convergence to the interlocutors is a quasi-automatic process, without any motivation concerning attitudes or identity. Trudgill's theory was opposed by many sociolinguists (e.g.,

see discussion in *Language in Society* 37(2) 2008; Kerswill, 2010); and, in addition, some recent experiments have also found evidence supporting the effects of attitudes in speech convergence (e.g., Babel, 2012; Pardo et al., 2012; Yu, Abrego-Collier, & Sonderegger, 2013). When this type of short-term convergence or divergence between speakers leads to long-term accommodation, as in some language/dialect contact situations, the role of attitudes is often discussed at the community level (e.g., Stausland Johnsen, 2015), while at the individual level, the connection between individuals' social attitudes and their linguistic behavior is much more complex and is less often studied. As Stausland Johnsen (2015) notes, we rarely have sufficient information about individual speakers' attitudes and social networks (p. 622). Moreover, the complex nature of attitudes also adds to the difficulty in exploring attitude-behavior correlations (Ladegaard, 2000). Gallois (2013) pointed out that most language attitude studies have placed a strong emphasis either on "language" (a more linguistic approach) or "attitudes" (a more socio-psychological approach). She encouraged future studies to combine different approaches and assess both language production and attitudes in the same research agenda (pp. 171-172).

In previous studies where attitudes were not the focus of the research but were brought up to interpret linguistic patterns, speakers' attitudinal information was either gleaned from sociolinguistic interviews (e.g., Llamas, 2007; Haug-Hilton, 2010) or measured in an arguably over simplistic way (e.g., Haddican, Foulkes, Hughes, & Richards, 2013); thus one cannot make very strong claims about the possible link between attitudes and behavior. Studies that concentrated on attitudes often employed various direct and indirect methods of measurement, but the measured attitudes were not consistently found to correlate statistically with linguistic behavior (e.g., Ladegaard, 2000; Stuart-Smith et al., 2013). Explicit attitudes and implicit attitudes were also found to behave differently in predicting linguistic behavior. Kristiansen (2009) argued for the role of subconsciously offered attitudes as a driving force in linguistic variation and change, but the consciously offered attitudes in his study were not in accordance with speakers' linguistic behavior. The study presented in this thesis collected speakers' consciously offered attitudes by employing a

newly developed attitudinal questionnaire methodology introduced by Llamas and Watt (2014), and aims to explore whether speakers' explicit attitudes could predict linguistic behavior¹.

Therefore, with respect to the exploration of the attitude-language correlation, I ask the following questions in this thesis:

1. In large-scale dialect contact situations, will speakers' linguistic variation be conditioned by their social attitudes? If so, how will speakers' attitudes influence their linguistic production?
2. Are speakers' explicit attitudes collected by overt questionnaires likely to predict linguistic behavior?

When investigating the attitude-language correlation, a closely related issue is how speakers' attitudes can be teased apart from their contact with other people. Trudgill's deterministic model argues that speech convergence largely depends on the amount of personal interaction speakers have, and rejects such social factors as attitudes and identity. By contrast, the CAT model advocated for the role of both attitudes and interaction. However, when empirically testing these models, a practical issue is raised due to the possible correlation between positive attitudes and more social interaction, as people are inclined to engage with those of whom they think favorably (e.g., Labov, 2001). Some recent small-scale studies have shown that attitudes are likely to play a role in speech convergence that is independent from the effects of social interaction (Pardo et al., 2012; Stuart-Smith et al., 2013). However, in larger-scale studies of dialect contact between speech communities, the situation of speakers' attitudes and social networks is often more complex. The specific questions I ask in this issue are:

¹ Subconsciously offered attitudes are not explored in this thesis.

3. Will speakers' attitudes play an independent role in language change if their social contact is also taken into consideration?
4. How will attitudes interact with social contact to influence speakers' linguistic production?

1.2 Attitudes in koinéization

This dissertation aims to explore the role of attitudes in a specific dialect-contact context: koinéization, where speakers of different but mutually intelligible varieties of the same language come into contact and a new mixed variety is formed (Siegel, 1985). The process of koinéization is a form of rapid and occasionally dramatic language change induced by dialect contact (Kerswill, 2002, p. 669), in which the newly forming variety (i.e. a “koiné”) is much less stable than a longer established dialect, and linguistic variation between individuals within the community is larger (Siegel, 2010a). Koinéization situations have been studied from various perspectives in many western societies, such as New Zealand (Trudgill, 2004), Bergen in Norway (Kerswill, 1994), the English Fens (Britain, 1997), Milton Keynes in England (Kerswill & Williams, 2000), and Høyanger in western Norway (Omdal, 1977; Trudgill, 1986, pp. 95-106; Solheim, 2009) (see section 2.2). Similar scenarios have also been studied in Chinese societies, such as the Kundulun District in Baotou (e.g., Xu, 1992; Zhu, 2010; Xu, 2010), the Jiangnan oilfield in Hubei (Sun, 2011; 2013), the Qingshan District in Wuhan (Lu, 2014) (see section 2.4). Some of these studies have explored the social and linguistic constraints on the variation in this sort of mixed variety using the variationist sociolinguistic approach, but the social factors discussed at length are often demographic factors as sex, age, social class, and region (e.g., Kerswill, 1994; Kerswill & Williams, 2000), whereas the role of speakers' attitudes has been less studied. This thesis presents the case of Hohhot, which is the provincial capital of Inner Mongolia in northern China. It is home to a complex mixture of ‘traditional’ local residents, who speak the local Jìn dialect (Jìn Yǔ

晋语), and migrants, who arrived in the 1950s and 1960s encouraged by government policy, speaking Pǔtōnghuà (普通话 ‘standard Mandarin’) with the accents and dialects of their original hometowns. Thus, a mixed, new vernacular combining features of Jìn and Pǔtōnghuà was formed, known locally as “Hū Pǔ (呼普)”, which means ‘Hohhot Pǔtōnghuà’. Anthropological studies of Hohhot have shown that there have been intense social conflicts between the locally-born and migrant communities (Jankowiak, 1993; Borchigud, 1996; see section 3.3.1). Therefore, this thesis explores how these attitudinal factors could influence Hohhot people’s speech production in different stages of the formation of Hū Pǔ, and more generally, investigates the role of socio-psychological factors in new dialect formation scenarios. The following questions are to be addressed concerning this issue:

5. What is the role of attitudes in koinéization? Will attitudes play different roles in different generations of speakers in a new dialect formation scenario?

6. Are social factors like speakers’ attitudes and identities important in shaping the outcome of koinéization?

In order to address these questions, I conducted the first large-scale variationist investigation of Hohhot, and interviewed 67 speakers from both the migrant and the locally-born communities. All speakers’ attitudes were quantitatively measured using an overt questionnaire, and their linguistic data were collected from interviews and an elicitation task (see chapter 4). The linguistic feature examined is a set of disyllabic words known as “l-words” (Hou, 1999) because the second syllable always has an “l” as the onset, such as /xuǎ⁴³la³⁵/ 划拉 ‘scribble’ (see section 4.4.1). L-words display variation at different linguistic levels, such as in the vowels, consonants, tones, and stress patterns. Among these features, this thesis focuses on two linguistic variables. The first is the stress pattern variable: l-words display variation between weak-strong and strong-weak stress patterns. A weak-strong stress is more commonly associated with Jìn dialect and a strong-weak pattern is typical of Pǔtōnghuà. The second is the fricative variable: for l-words with the initial sounds of [p^h, t^h, k^h, h], the sounds are likely to be produced as [p^x, t^x, k^x, x], with a period of

frication, typically [x]. This study explores whether speakers of Hū Pǔ vary in the degree to which they adopt these two features, and whether this variation is conditioned by their attitudes.

1.3 Attitudes and awareness

Another topic to be addressed in this dissertation is the relationship between attitudes and the level of awareness of the linguistic feature. This issue actually appeared later in the data analysis phase of the study. The fricative variable was not selected as a target linguistic feature at the beginning of the research because I – even though a native speaker of this variety – had no awareness of it, but it was found to be of potential interest when I was analyzing the participants’ speech in their interviews. Evidence from the interviews with the speakers suggested that the fricative variable and the stress pattern variable were subject to very different levels of awareness. The weak-strong pattern was a salient Jìn dialect feature and attracted speakers’ overt social commentary in the interviews, but for the fricative variable, speakers did not show explicit awareness of it.

In sociolinguistics, the link between speakers’ awareness of a linguistic feature and their behavior with respect to the feature has long been discussed. In dialect contact situations, it is traditionally considered that explicit awareness or the salience of a linguistic feature is a “necessary but insufficient condition” for the acquisition of new dialect features (Auer et al., 1998, p. 184). That is, for a feature to be acquired, it must be salient enough to be noticed by speakers (e.g., Siegel, 2010b). This condition is also emphasized when exploring the attitude-language correlation. As Labov (1963) put forward, “we would like the feature to be salient, for us as well as the speaker, in order to study the direct relations of social attitudes and linguistic behavior” (p. 8). However, recent work by Nycz (2016) has found that speakers acquire new dialect features even though they are not explicitly aware of them. She thus claimed that explicit awareness is not a prerequisite for individuals’ dialect change.

Moreover, evidence from Drager and Kirtley (2016) suggests that speakers' linguistic production is likely to vary according to their positive or negative attitudes even when they have no explicit awareness of the linguistic features. Therefore, the present study further explores whether an attitude-language correlation could be found in linguistic features that are below the level of speakers' conscious awareness, that is, the fricative variable in this research; and I ask:

7. How will the effects of attitudes be related to the level of awareness of the linguistic variable itself? Is explicit awareness a threshold for attitudes to play a role in linguistic production?

1.4 Thesis structure

This thesis comprises eight chapters. Chapter 2 reviews major discussions related to attitudes in the previous literature, focusing on the challenges in the exploration of the attitude-language correlation, and the effects of attitudes on speech convergence in relation to speakers' social contact and their level of awareness of the linguistic feature. The role of attitudes in koinéization is also discussed by reviewing previous studies of new dialect formation scenarios in both western localities and Chinese societies.

In chapter 3, the social and linguistic background of the locality of Hohhot is provided. I present how the new variety of Hū Pǔ emerged in the intense social conflicts between the migrant and locally-born communities, and why Hohhot is a fruitful site for variationist studies.

Then in chapter 4, I present the methodology and elaborate on the linguistic and social variables to be used in the analysis, with a special focus on how speakers' attitudinal information and their social network data were collected and quantified. I also introduce

how the word elicitation task was designed and successfully used to collect l-words data, and how the linguistic variables were selected.

Chapters 5 and 6 present a statistical approach to the correlation between the social and linguistic variables, which is supplemented by qualitative discussions with evidence from the sociolinguistic interviews. Chapter 5 focuses on the results of the stress pattern variable. I explore the social constraints operating on speakers' use of l-words stress patterns, with a particular interest in the effects of attitudes and their interaction with other social factors like age, sex and social networks. Chapter 6 reports the fricative variable results, as well as the issues investigated in the stress pattern variable, and it also focuses on how the effect of attitudes is related to speakers' explicit awareness of the linguistic feature.

In chapter 7, I summarize and discuss the results of the two linguistic variables collectively, and return to the issues raised in chapter 2 and in the research questions, exploring the connections between language use and social attitudes in large-scale dialect-contact contexts.

The final chapter presents the conclusion and implications of this thesis, and outlines some limitations and future directions of the research.

2 Background

This chapter discusses the existing literature on the main issues of this dissertation. Section 2.1 reviews the main challenges in investigating the attitude-language correlation including the complex nature of attitudes and how to tease apart attitude from social interaction. Section 2.2 focuses on previous discussions of the role of attitudes in new dialect formation or koinéization contexts. Section 2.3 elaborates on the notion of social awareness and salience in relation to the effects of attitudes. The discussions in the above three sections will be mostly based on studies conducted in western societies, and section 2.4 will especially focus on relevant research in the Chinese context.

2.1 Attitude-language correlation

Although speaker attitudes have been often studied in sociolinguistics, there is a consensus that it is “no simple undertaking” (Llamas & Watt, 2014, p. 616). Sociolinguists often discuss the effects of speakers’ attitudes on language change, but various problems occur when attempts are made to accurately verify this correlation. This section will review two major issues related to the study of the relationship between attitudes and language use. The first issue lies in the complex nature of attitude itself and the difficulty in studying explicit attitudes as predictors of linguistic behavior. The second issue relates to the challenge of teasing apart attitude from the effect of social interaction. Some previous small-scale studies have shown that the effects of attitude and interaction might be independent (see

discussions in section 2.1.2), but the current project is tackling the problem in a larger-scale dialect contact situation.

2.1.1 Attitude and its complicated nature

The term *attitude* is defined as a predisposition to evaluate some object with some degree of favorableness or unfavorableness (Eagly & Chaiken, 1993; Ajzen & Cote, 2008), or simply as “an evaluation of an object of thought” (Bohner & Wanke, 2002, p. 392). Scholars usually have disagreements about the definitions of attitudes from different angles in different theories (Mckenzie, 2010). Real-life attitudes are usually multifaceted and multilayered: they cannot be directly observed and need to be inferred from individuals’ responses to the attitude object (Schwarz, 2008). Thus they are difficult for researchers to capture or measure. The topic of correlation between attitudes and behavior has been of particular interest to social psychologists (see e.g., Alwin, 1973; Mostyn, 1978; Glasman & Albarracin, 2006). However, as Ajzen and Cote (2008) noted, although attitudes can help us understand general patterns of behavior, they are usually poor predictors of specific behaviors with respect to the object of the attitudes (p. 305). The distinction between explicit and implicit attitudes is also emphasized when exploring the attitude-behavior relationship. Explicit attitudes, or overt attitudes are often measured with direct self-reported evaluations and have long been used by researchers, whereas implicit attitudes, or covert attitudes have been increasingly studied in recent decades due to the development of various new measurements. Bohner and Dickel (2011) proposed that measures of implicit attitude usually predict spontaneous, less controllable behavior, while explicit attitudes predict deliberative, more controllable behavior. In the field of sociolinguistics, the connection between speakers’ attitudes and their linguistic behavior is also a frequently studied topic. However, due to the complicated nature of attitude, the correlation is often difficult to test.

Many sociolinguistic studies use qualitative attitudinal information gleaned from interviews to explain speakers’ linguistic behavior, claiming a possible correlation between attitudes and language change. For example, Llamas (2007) studied an urban variety of

British English in Middlesbrough, which is situated in a border area in the north of England between “the extreme south of the North East and the extreme north of Yorkshire” (p. 580). The transitional nature of Middlesbrough resulted in complicated regional identity construction for the local speakers. Llamas examined local speakers’ use of the glottalized voiceless stops (p t k) and collected their attitudes towards their language and area using an identification questionnaire composed of seven general questions. Through a qualitative discussion, Llamas found that speakers’ linguistic convergence with North Eastern variants and divergence from Yorkshire forms were correlated with their attitudinal information, which outlined their shifting sense of regional identity. Haug-Hilton (2010) followed the methodology of Llamas (2007) in eliciting direct attitudes from speakers in her investigation of Hønefoss, Norway, and expanded the identity questionnaire by adding more indirect questions about informants’ linguistic and regional identity. She then qualitatively discussed the potential link between speakers’ socio-psychological orientations and the variation in their linguistic production, but noted that the data set was not large enough to investigate this correlation in depth. Another study by Clark and Watson (2016) examined the phonological leveling and diffusion of the realization of /t/ as [h] in Liverpool and two localities in its hinterland (Skelmersdale and St. Helens). In the investigation of Skelmersdale, they found that the constraints operating on this linguistic feature were not the same for all speakers, and they attributed this difference to speakers’ positive or negative attitudes towards Skelmersdale and Liverpool by providing attitudinal evidence from sociolinguistic interviews. Therefore, it is very common in sociolinguistic studies to bring up this sort of qualitative discussion about speaker attitude as a potential factor to interpret linguistic changes. However, it is difficult to make a strong argument about the effects of attitudes in this context. As Clark and Watson (2016) put forward, using selected extracts from interviews might run the risk of being accused of ‘cherry picking’ (p. 56).

Therefore, more attempts have been made to quantify the attitude data collected in interviews in order to test the correlation between attitudes and language use. However, in

this type of study, the attitude measurement was usually too simplistic to draw firm conclusions, probably because the role of attitudes was often not the focus of the study. For example, Haddican et al. (2013) in their study of the Northern English dialect of York, found a significant correlation between the speakers' degree of adopting FACE/GOAT diphthongization and their identification with the local community. Speakers' attitudinal information was measured in a simple manner by adding up the scores given to each speaker's answers to four questions related to their attitudes and local identity. Røyneland (2005, as cited in Haug-Hilton, 2010) investigated language variation and change in two towns in East Norway: Røros and Tynset. The study employed an explorative statistical method "correspondence analysis" to visualize the relationship between 15 linguistic variables and 12 social factors including speakers' attitudinal orientations. The results showed that there was a clear correlation between positive attitudes towards the local dialect/town and speakers' use of local linguistic features. However, although Røyneland considered several aspects of the speakers' attitudes and identity, the measurement was only in the form of a three-level gradient of positive, neutral and negative.

Other studies that have focused on attitudes have adopted various measurement methods to collect and quantify speakers' explicit and implicit attitudes, such as the use of overt questionnaires, the Matched Guise Technique (MGT) or Verbal Guise Technique (VGT) (see an overview of language attitudes measurement studies in, Baker, 1992, pp. 17-20; Watson & Clark, 2015). However, the quantitative measures of attitudes were not consistently found to be predictors of linguistic behavior. Ladegaard (2000), for example, assessed individual speakers' attitudes using both a verbal guise experiment and a questionnaire; however, the quantitative attitudinal data obtained from the VGT experiment were found to have no significant correlations with speakers' linguistic behavior. He then argued for the role of attitudes qualitatively using the information gleaned from the questionnaire, proposing that eclectic approaches should be used in future research concerning the correlation between attitude and language behavior. Ladegaard and Sachdev (2006) investigated 96 Danish learners of EFL about their attitudes towards British and American English using the

verbal-guise experiment, and found that although the vitality of American culture was acknowledged, the participants had no desire to adopt an American accent. Stuart-Smith et al. (2013) looked at the speech of adolescents in inner-city Glasgow, and found correlations between their linguistic convergence to the London accent with some social practices and their psychological engagement with a London-based soap opera. Speakers' attitudes towards accents in different areas were also measured in a perceptual test in this study, but little correlation was found between the attitudinal factors and linguistic production. Stuart-Smith et al. (2013) claimed that this could be because the attitudes were measured in an explicit way, and they suggested that future research to elicit covert attitudes should be "via questions not specifically about language/dialect, but strongly associated concepts" (p. 528). Kristiansen (2009) emphasized the distinction between conscious and subconscious attitudes. As part of the LANCHART (language change in real time) language attitude project in Denmark, Kristiansen (2009) investigated groups of adolescents' evaluations of stimulus speakers with various Danish accents, and the data collection contexts allowed a clear distinction between overtly and covertly offered evaluations/attitudes. He found that the informants from different regions usually identified their local variety as their preferred Danish speech style when they provided conscious evaluations. However, in the non-conscious mode, they preferred the modern Copenhagen variety, which was also found to be the most influential variety in the entire country. This study was later replicated in seven other Nordic communities, and similar results were found (Kristiansen, 2015). Therefore, Kristiansen argues that subconsciously offered attitudes appear to be a driving force in linguistic variation and change in a way that consciously offered attitudes are not. Recent work by Stausland Johnsen (2015), through a meta-analysis of the previous sociolinguistic studies in Norway, advocated for the role of attitudes in linguistic accommodation and language diffusion, proposing that we should investigate speakers' attitudes when explaining observed diffusion patterns.

To summarize, although many studies have claimed that there is a relationship between speakers' attitudes and their linguistic behavior, it is rare for a statistically significant

correlation to be found, especially in the case of explicit attitudes consciously offered by speakers. Recent work by Llamas and Watt (2014) proposed that the technique of the “Attitude Analog Scale”, as a modified version of magnitude estimation (Redinger, 2010), could be innovatively utilized in sociophonetic studies. This type of written questionnaire offers a quick and valid method to collect direct and quantitative attitudinal data in large-scale fieldwork (see section 4.5); thus it may provide a new route to the investigation of the attitude-language correlation. This dissertation applies this methodology to the Hohhot context, and explores whether the direct, overt attitudes obtained from questionnaires can predict linguistic behavior.

2.1.2 Attitudes and social interaction

Another challenge in studying the relationship between attitudes and language use is teasing apart speakers’ attitudes from their social networks. In Trudgill’s “deterministic model”, or what Stausland Johnsen (2015) calls the “interaction-only model”, speakers’ frequency of contact with a certain dialect or community was considered to be the most important factor in linguistic accommodation, which was described as merely a matter of “who interacts most often with whom” (Trudgill, 2008, p. 251). Attitudes and identity factors, in this process, were taken to be irrelevant. By contrast, the CAT model advocates for the effects of both attitudes and interaction, predicting that speech convergence depends on not only speakers’ amount of personal interaction, but also what kind of attitudes they have towards each other. However, when empirically investigating the main tenets of this model, it is difficult to tease apart the effect of attitudes from social interaction. It is very likely that people who have more positive attitudes towards speakers of a certain community would be more inclined to make friends with them and integrate into this group. Therefore, even if effects of attitudes are found, it is difficult to know whether this is because the speakers have more positive attitudes to a certain group or whether they simply have more contact with the group.

Some previous work has attempted to explore the role of attitudes and interaction in the same study. For example, in Stuart-Smith et al. (2013), as previously mentioned, speakers' psychological engagement with soap opera characters was found to correlate with their speech, but at the same time, their reported exposure to television was not a significant predictor. Pardo et al. (2012) examined how roommates in college converge to each other's speech, and speakers' self-reported feelings of closeness were found to play a role, while their relationship quality, such as how many hours they spent together or how many meals they had together, did not correlate with their degree of speech convergence. These studies have shown that speaker attitudes are likely to play a role in speech accommodation that is arguably independent from, or at least is not exactly the same as, the effects of interaction. However, in larger-scale studies of speech communities, the descriptions of attitudes and social networks are usually much more complicated. Khan (2006), for instance, investigated a multi-ethnic community in Birmingham, England, and explored how social network and social psychological orientation could influence linguistic behavior using a quantitative variationist approach. The findings were different in the three heritage groups he examined. For Pakistani and Afro-Caribbean speakers, their ethnic identity was a stronger predictor of the phonological variation than either social network or attitude towards Birmingham, whereas for Anglos, social network could better predict their use of ethnic variants. Therefore, it is essential to consider this issue in specific social and linguistic environments. As I explain below, the current Hohhot project quantitatively measures speakers' social interaction with local Jin speakers, and explores its effect in the same statistical model with speakers' attitudes towards the local dialect and community. The findings will provide more evidence for the interaction of attitudes and social networks in larger-scale dialect contact contexts.

2.2 Attitudes in koinéization

This thesis explores the role of attitudes in koinéization situation where rapid linguistic change is induced by contact between mutually intelligible dialects, and new, distinctive varieties are formed. The language change situation is often the result of human migration, such as in colonial settings, new towns and rapid urbanization, where new speech communities are formed through the influx of a large number of newcomers or the transplantation of people from different speech communities to a new location (Kerswill, 2010). The early stage of the process, according to Trudgill (1986), is usually dialect mixing which involves enormous amounts of linguistic variability. However, the number of variants available will be gradually reduced in later stages, as a result of accommodation between speakers through face-to-face interaction. This reduction is composed of processes like *leveling*, in which linguistically marked and demographically minority speech forms are eradicated, and *simplification*, which increases the grammatical regularity and reduces formal complexity. Sometimes the minority forms also survive through *reallocation* where they are refunctionalized, evolving new social or linguistic functions in the new dialect (Britain & Trudgill, 1999). As the outcome of these processes, the new variety, in most cases, will acquire norms and a certain stability, which is defined by Le Page and Tabouret-Kelly (1985) as *focusing*.

This sort of dialect mixing and new dialect formation process (Trudgill, 2004) has been discussed at length in many different societies all over the world, and it seems to have two types of scenarios (Kerswill & Trudgill, 2005; Kerswill, 2010). One scenario is similar to that which Trudgill referred to as a *tabula rasa* situation, in which the settlement is in a relatively large territory and there is no antecedent population speaking the same or a related language (Trudgill, 2004), such as the formation of New Zealand English (Gordon et al., 2004; Trudgill, 2004), Hindi in South Africa (Mesthrie, 1992) and Fiji (Siegel, 1987; 1997). The second scenario of new dialect formation is that of the new town, where the settlement is in a “geographically delimited area” (Kerswill & Trudgill, 2005, p. 196) and there are often

prior speakers of the same language, so intense interpersonal communication could take place between the existing speakers and the newcomers. A number of typical new town situations that have been studied include the English Fens (Britain, 1997), Milton Keynes in England (Kerswill & Williams, 2000), Høyanger in western Norway (Omdal, 1977; Trudgill, 1986, pp. 95-106; Solheim, 2009; Kerswill, 2010, pp. 240-242), among others. These studies have discussed the different stages involved in koinéization, and show that the focusing process in these contexts often takes two or three generations. Trudgill (2004) describes this new dialect formation process using a three-stage model on the basis of his study of the New Zealand English. In Stage I, the speech of the original adult immigrants goes through rudimentary leveling through accommodation. The second-generation children, in Stage II, will select variants from different dialects “at will” from a supermarket-like linguistic feature pool, and Stage III is the formation of the focused or crystallized new dialect. The first native-born generation, or the children of the original immigrants, is normally considered as the crucial stage in new dialect formation (Kerswill & Trudgill, 2005).

Previous studies also tried to explore possible factors that could influence the outcome of the dialect mixture. The role of demographic factors has been discussed at length from various aspects. For example, the proportions of speakers of particular dialects in the original contacting groups, and the relative frequency of certain dialect variants were often the main factors brought up to explain the realization of the new variety (see e.g., Trudgill, 1986; Siegel, 1997). Kerswill and Williams (2000, p. 75) also listed other potential factors such as the proportion of children to adults in the immediate post-settlement years, and the degree of linguistic difference between the contributing varieties. Britain (1997), through his investigation of the koinéization process in the English Fens, found different rates of focusing for two linguistic features: one of them focused relatively quickly by the reallocation of different variants, but the other one formed an unstable situation and was much slower in crystallizing a distinct form. He attributed this distinction to the relative salience and complexity of the features, and moreover, the social structure in which no universal schooling environment was present in the sparsely populated Fens further

impeded the focusing process, because children did not have a stable target adult model or peer group to conform to. A similar situation was found in Mæhlum's (1992) study on the Norwegian Arctic territory of Spitsbergen. The second generation children there retained a strong influence from their parents' speech, because the migrant families frequently spent long summer breaks in their place of origin on the Norwegian mainland. Sudbury's (2000, 2001) study on the Falkland Island English presented another interesting demographic factor. Although the dialect contact situation began from the mid-19th century, Falklands English was still unstable with considerable within-group and within-individual variation. Apart from the isolation of communities in the Falklands, Sudbury also pointed to the transience of the population as a potential factor, because many of the settlers were contract workers, resulting in high rates of in- and out-migration (Sudbury, 2000, as cited in Kerswill & Trudgill, 2005).

However, when it comes to the way socio-psychological factors could influence the outcome of new dialect formation, such as speakers' identities and attitudes, the situation is usually more complicated and it is more difficult to test their relationship with linguistic change. I have mentioned Trudgill's (2004, 2008) deterministic model when discussing the interaction between attitudes and social networks in section 2.1.2 (see also section 1.1), in which Trudgill argues for the quasi-automatic accommodation between speakers and rejects the role of social factors. This model was brought up on the basis of Trudgill's analysis of the formation of New Zealand English and several European languages in colonial settings, which were closer to tabula rasa scenarios. He claimed that the outcome of new dialect formation was predictable if we knew the geographical and demographical backgrounds of the population. Social factors like prestige and identity counted for nothing in Stages II and III, he argues. Children of the original migrants were selecting features at random from the feature pool because they did not have ideas about prestige or identity-marking functions in the tabula rasa context (Trudgill, 2004). Trudgill's argument on the irrelevance of identity has received numerous criticisms from other sociolinguists (see discussion in e.g., *Language in Society* 37(2) 2008; Kerswill, 2010; Solheim, 2009). The major

counter-arguments are from two perspectives. First, Trudgill's deterministic model is restricted to *tabula rasa* situations, which is more like an ideal experimental situation and is rarely found. Even when there are situations like this, the *tabula rasa* "quickly ceases to be a clean slate" because adults in the new community are not entirely deracinated (Kerswill, 2010, p. 238). Therefore, as Solheim (2009) argues, it is problematic to talk about *tabula rasa* societies at all (p. 204). Another criticism is about Trudgill's narrowing down the concept of social factors to identity or even national identity. As Holmes and Kerswill (2008) put forward, to imply that "national identity can stand for all types of identity deflects our attention from the real sociolinguistic issues" (p. 274). Bauer (2008) also mentions that the social factors involved will not be as simple as feeling that one is 'British' or 'New Zealand'; it will be much more local and much more specific (p. 273). Kerswill (2010) proposes that even if prestige or identity is ruled out at the very start of the *koinéization* process, other factors such as "parents' inherited ideas about good and bad behavior, and acceptable linguistic practices including politeness" (p. 238) would still play a part. To conclude from previous discussions, it is essential to study new dialect formation in specific local conditions, especially when complex social factors are taken into account. The locality of Hohhot in the current thesis is not a *tabula rasa* context. It is closer to the new town scenarios where there are prior speakers of the same language. The social conflicts between the locally-born residents and the in-migrants add to the complexity of the social situation. Therefore, it is interesting to explore whether later generations' choices of linguistic features will be affected by their social-psychological orientations. As Kerswill and Williams (2000) notes, sociolinguists have realized that in order to understand behavior that leads to the adoption or rejection of potential changes, we need to go to the individual speakers (p. 65). Schneider (2008) also proposed that it would be "a worthwhile task" to design a study that tests "the straight-forward connections" between socio-psychological attitudes and the use of specific linguistic forms (p. 266). In this context, the variationist sociolinguistic approach is often used to quantitatively explore the relationship between speakers' adoption of certain linguistic features and their social characteristics (e.g., Kerswill, 1994; Kerswill & Williams, 2000). However, although such traditional demographic factors as sex,

age, social class and region have been frequently investigated, the role of speakers' attitudes and identity has been less often systematically studied. The factor of social networks is also brought up in some studies to explain speakers' linguistic behavior; however, few studies have examined its effect in relation to attitudes in koinéization. In the remaining part of this section, I review some discussions on the role of social-psychological factors and social networks in new dialect formation scenarios.

Kerswill and Williams (2000) investigated a forming koiné in the English new town of Milton Keynes, which was designated in 1967 and received a large number of migrants mainly from southeast England in the following decades. Unlike other koinéization studies which were mostly post-hoc observation of completed changes, Kerswill and Williams concentrated on the process of the koiné formation. They recorded 48 young children (ages 4, 8 and 12) and their principal caregiver (in most case the mother), and quantitatively examined ten phonetic variables in their speech, claiming a crucial role of demographic factors for the outcomes of koinéization. For example, the features found in the majority of the input varieties often win out in the Milton Keynes koiné. Complexity of features also has some effect where phonologically and lexically simple features are more often adopted. Additionally, they also examined speakers' social network characteristics, and found that children who were well integrated into their peer group had a higher degree of fronting in their GOAT vowel. Kerswill and Williams (2000) thus claimed that, in the formation of the Milton Keynes koiné, an individual speaker's adoption or rejection of features depended on his/her social networks. However, in this study, the discussion of the social networks factor was qualitative, and psycho-social factors like attitudes were not further explored.

Another koinéization study by Kerswill (1994) took into account both social networks and attitudinal factors. He examined the speech of the in-migrant community from rural Stril to the city of Bergen, Norway, and explored how individual speakers' social characteristics could influence their adoption of three Bergen dialect features (morpho-lexical features, schwa-lowering, and tonemicity). The demographic factors examined in this study were measured in a number of different dimensions, such as social class, which was described by

family background, occupation, and ownership of property. Kerswill also assessed migrants' social networks in terms of the 'Strilness' or 'Bergenness' of their contacts, which were measured from a number of different dimensions like contacts with the rural district, origin of household members, membership of a rural migrant association, contacts at work, and contacts outside of work. Moreover, Kerswill also measured speakers' attitudes towards the Stril dialect by simply assigning a score ranging from 0 to 3 to each speaker (representing a gradient from negative to increasingly positive attitudes). Quantitative analysis found that the attitudinal scores were significantly correlated with migrants' use of Bergen morpho-lexical features, and social network was a significant predictor for the schwa-lowering feature. However, Kerswill claimed that these results were not quite reliable due to potential intercorrelations between the social factors. When he tested the independent effects of these factors later, attitudes and social networks became insignificant. He then discussed the role of social networks and social-psychological factors in individuals' use of morpho-lexical features in a qualitative analysis. The Bergen in-migrant study by Kerswill (1994) is an important one that pays attention to the effects of speakers' attitudes and social networks at the individual level, and quantitatively investigated their relationship with linguistic production. However, the attitudinal factor was only restricted to language attitude towards Stril and was measured in a simplistic way. The limitations addressed in the statistical analysis also made the arguments less convincing.

Other new town scenarios of new dialect formation were also found on the shores of fjords in western Norway. Odda and Tyssedal, for example, were two small towns which emerged in the early twentieth century because of the development of the ore-smelting industry. The two towns were only five kilometers apart, but formed radically different koinés, reflecting the dialects spoken by the in-migrants of different geographical origins (Sandve, 1976; Kerswill, 2002). The Odda koiné was closer to the western Norway dialects because the majority of the population of the original migrants was from the west. By contrast, the Tyssedal koiné was found to resemble the eastern varieties, especially in morpho-lexical features, despite the fact that only 35% of the in-migrants were from eastern Norway. This

was because the linguistic situation in Tyssedal was more complex and diverse with no single dominating dialect group. Speakers of various eastern dialects still formed the largest group, and additionally, the eastern dialect features were closer to the standard forms, thus they were more likely to survive in this sort of heterogeneity context (Kerswill, 2002).

Another western Norwegian koiné, often referred to as the classic new dialect formation situation, was formed in the new town of Høyanger. The social and linguistic context is similar to that of Odda, with the majority of incomers from western Norway. The earlier linguistic situation was reported by Omdal (1977) in a brief paper using dialectological methods. Trudgill (1986) then discussed the Høyanger data as a typical new dialect formation example in his work on dialect contact. Later, Solheim (2009) re-examined data from several older studies (collected in 1956 and 1975), and conducted new fieldwork in Høyanger in 2001, which made it possible for her to study the stages of dialect development in Høyanger from both a diachronic and a synchronic perspective. As Kerswill (2010) noted, the social influence on the outcome of new dialect formation was discussed extensively in the Høyanger study. Due to the geographical origin of the in-migrants, the Høyanger koiné, unsurprisingly, included features from both western and eastern dialects. However, a large number of high-frequency words were found to have an eastern form, which was disproportionate to the number of East Norwegian migrants, comprising only 17% of the population. Solheim (2009) claimed that this was because of the high social status and prestige of the standardized eastern forms. However, some of the features that initially took the eastern forms were later replaced by the original western variants, because these forms, according to Solheim, were socially and geographically salient as markers of the standardized eastern varieties; thus they were too strong to be acceptable in the Høyanger koiné (see also Kerswill, 2010). The emerging of local community identity was also found in Høyanger. An interesting example given by Solheim was the pronunciation of the place name “Høyanger”. She found that, although the use of traditional diphthong [ɔy:] was declining in general among speakers of recent generations, they still chose to pronounce the name of Høyanger with the traditional diphthong quality, which was considered to be a

strong symbolic marker of local identity. Solheim (2006, as cited in Kerswill, 2010, p. 242) argued that, from her observation and communication with the second generation speakers, individuals' language use was to a large extent influenced by their personalities and life worlds. The findings of the Høyanger study support the role of social factors as motivation for linguistic change in koinéization, and provide counter-arguments to Trudgill's deterministic model. As Solheim (2009) put forward, migrants will always bring with them the values and norms of their families and places of origin, so it is controversial to study new dialect formation without considering social or individual factors (p. 204).

2.3 Attitude, awareness and salience

Another issue brought up in this dissertation is the effects of attitudes in relation to speakers' explicit awareness of the linguistic feature, or the salience of the feature. In sociolinguistics, linguistic changes are classified by Labov (1972) into *change from above* and *change from below* according to whether speakers are consciously aware of them. This distinction between consciousness and unconsciousness has been found to be ubiquitous in Labov's framework of language variation and change, such as overt and covert social values, as well as careful and casual speech styles (Kristiansen, 2011). At the level of linguistic features, Labov (1972) classifies linguistic variables into *indicators*, *markers* and *stereotypes* in terms of social awareness or salience. Indicators are never the topic of social comment and are below the level of conscious awareness with no stylistic variation, while markers are more likely to be noticed by listeners and show consistent stylistic and social stratification. Stereotypes are the most salient variables, which will often attract overt social commentary, especially to the stigmatized variant. Another similar model proposed by Silverstein (2003) is the "order of indexicality", in which higher orders of indexicality are increasingly salient, and the first-, second-, and third-order indexicality could roughly correspond to Labov's indicators, markers and stereotypes respectively. However, Silverstein's indexicality emphasizes more on how linguistic forms acquire social meaning

(Rácz, 2013). On the basis of these concepts, Johnstone, Andrus, and Danielson (2006) investigated the enregisterment of a dialect in Pittsburgh called “Pittsburghese”, and described how previously unnoticed features of “first-order indexicality” could rise to be used as “second-order” markers of social class, correctness or localness, and later are used by speakers as “third-order” indexicals to perform their self-conscious local identity. They explored different indexical meanings associated with the local linguistic feature monophthongal /aw/ reported by Pittsburghers. First-order indexicality was seen in nonmobile speakers who were not aware of the correlation between this feature and geographical location and insisted that “everybody spoke [speaks] that way” (Johnstone et al. 2006, p. 82). Second-order indexicality arose when speakers started to associate these variants with meanings like working class or incorrectness and shift styles in their own speech. Third-order indexicality was found when speakers explicitly talked about this stylistic variation, and both Pittsburghers and non-Pittsburghers could use this variable to perform local identity (often in ironic ways).

So why are some linguistic features more likely to arise as markers or even stereotypes than others? Or in other words, why are some features more salient than others? Salience is “a property of a linguistic item or feature that makes it in some way perceptually and cognitively prominent” (Kerswill & Williams, 2002, p. 81). It is often believed that the more salient a feature is, the more likely it is for an individual to attend to it and become aware of it (Drager & Kirtley, 2016). However, salience is a very complex and controversial concept in linguistics and is defined differently by different researchers (see discussions in e.g., Auer et al., 1998; Kerswill & Williams, 2002; Hickey, 2000). The salience of a linguistic feature is likely to be increased by a number of factors, such as its prosodic prominence (Yaeger-Dror, 1993), acoustic extremes (Podesva, 2006), frequency of occurrence (Bardovi-Harlig, 1987), and “localisedness” (Honeybone & Watson, 2013), among others. Trudgill (1986) comes up with four factors that could contribute to the salience of a linguistic feature: 1) being overtly stigmatized, and often because there is a high-status variant which tallies with the orthography; 2) being involved in linguistic change; 3) having variants that are phonetically

radically different; 4) being involved in the maintenance of phonological contrasts (p. 11). However, Trudgill's criterion has received some criticism from other researchers. Kerswill and Williams (2002) claim that although Trudgill's model is testable, some points in the criterion can be accused of circularity. For example, if salience is used as part of an explanation for language change, the change cannot be maintained as a condition for salience (p. 90). Kerswill and Williams conclude that salience should be operationalized with consideration of three components: the presence of a linguistic phenomenon whose explanation we suspect may be due to the salience of the linguistic features involved, language-internal factors such as the presence of phonological contrast or great phonetic distance, and language-external factors including e.g., cognitive, pragmatic or social psychological aspects. They further suggest that in order to give a circularity-free explanation to salience, the language-external factors are of crucial importance. More criticism about Trudgill's model was brought up by Honeybone and Watson (2013) concerning the vagueness of concepts in the criterion. They also advocate that salience is a gradient and should not be treated as an "all or nothing" matter (p. 311), claiming that the picture of salience of a particular linguistic variable has been masked in works that only look at one linguistic variable at a time, and future research should examine a number of linguistic features simultaneously to understand their relative salience. The complex nature of salience is further addressed in Rácz (2013), who suggests that the interpretation of the social meaning of markers is also likely to be dynamic, with two interesting examples: 1) Campbell-Kibler's (2011) study found that some variants in gay speech in American English could invoke particular stereotype personas, but they were only interpreted as such when these features were in conjunction with certain other variants; 2) Kristiansen, Maegaard, and Pharao (2011) in their study of Copenhagen Danish /s/, found that participants' interpretation of the fronted variants of /s/ strongly depended on whether the speaker was assumed to be an upper-middle class speaker living in the rich part of the town, or an immigrant living in another part.

Another indication from previous discussions is about the connection between speakers' awareness of a linguistic feature and their behavior related to that feature. As mentioned in section 1.3, it is often considered that the linguistic features that are being acquired or dropped by speakers in dialect contact situations must be salient enough to be noticed (see e.g., Auer et al., 1998; Watt, 1998; Siegel, 2010). However, this claim is challenged by the recent work of Nycz (2016). She interviewed 17 native speakers of Canadian English who had moved as adults to the New York City region, and examined to what extent their use of two dialect features had changed. The two features selected, Canadian Raising in /aʊ/ and the use of a contrast between the low back vowel /ɑ/ and /ɔ/, were subject to markedly different levels of awareness among speakers. The former was a stereotype of Canadian English which was constantly mentioned in metalinguistic commentary, whereas the latter was often below the level of conscious awareness. The results showed that the participants accommodated to the new dialect features in both cases. Out of the 11 speakers who showed evidence of having acquired the low back vowel contrast, six indicated no or limited explicit awareness of this feature. Nycz thus argues that explicit awareness is not a prerequisite for individual dialect change, and speakers' acquisition of linguistic features could happen unconsciously, driven by automatic accommodative processes. However, Nycz also suggests that the accommodative processes could be affected by speakers' explicit awareness of the variable and its social meaning, but the convergence of the feature that is subject to a low level of awareness will proceed unimpeded (p. 77). This may indicate that in the convergence process, attitudinal factors or evaluations associated with a linguistic feature are only likely to intervene if speakers are explicitly aware of the feature and its meaning, whereas for features below the level of awareness, the accommodative process will not be affected by attitude.

This actually highlights another issue of interest in the current dissertation, which connects to the interplay between attitudes and awareness/salience. Since explicit awareness is not a threshold for the acquisition of a linguistic feature, will it be a threshold for attitudes to play a role in speech production? That is to say, can we take it for granted

that speakers have to “know” about the variable and its meaning so that their attitudes can enhance or attenuate adoption of the feature? Some evidence to the contrary may be found in Drager’s ethnographic study of a girls’ high school (Selwyn Girls’ High) in New Zealand (Drager, 2009; 2011; Drager & Hay, 2012). Different social cliques were identified in this high school. The girl cliques who ate lunch in the Common Room set the norms of dress and behavior in the school, while other non-Common Room groups rejected these norms. Drager (2009) investigated the phonetic production of the quotative *like* in these girls, and found that Common Room girls produced longer /l/ durations and more monophthongal vowels. She claimed that, although the girls were not explicitly aware of these differences, the differences arose as a result of identity construction. Drager and Kirtley (2016) further discussed two speakers from two different non-Common Room groups, who demonstrated completely different attitudes towards the Common Room girls. The girl with positive attitudes produced realizations of quotative *like* that were similar to the Common Room girls’ speech, whereas the production of the girl with negative attitudes was realized as very different from the Common Room norm. Therefore, Drager and Kirtley claimed that the difference between the two girls’ production arose without any need for awareness of the sociolinguistic variable. However, this argument was only based on discussion of two individuals’ data, and their attitudes were simply defined as positive or negative.

In the case of Hohhot, the two linguistic variables embedded in l-words are also greatly different in levels of awareness among local speakers, which allows me to investigate the relationship between the relative salience or awareness of the features and the linguistic changes in which they are involved. While the correlation between attitudes and language is the main focus of this dissertation, this correlation is also tested separately in these two linguistic variables that are subject to different levels of awareness. Therefore, the findings will shed light on the question of whether the attitude-language correlation could be found in linguistic features below the level of conscious awareness.

2.4 Relevant studies in the Chinese contexts

The language situation in urban cities in China has experienced rapid change since industrialization began in the 1950s. A number of new industrial cities arose in the previously undeveloped areas where the population was boosted from a few thousand to half a million or more. The large-scale migratory movement was promoted by the government, and most immigrants were cadres, intellectuals, technicians and skilled workers, who were selected because they were most needed in the destination areas (Davin, 1999). As a result, immigrants from across the country with different linguistic backgrounds merged into the same area, leading to various language change and dialect contact situations. Apart from the immigrants, the local residents of the destination areas were usually the local dialect speakers, which added to the complexity of the linguistic situation, but provided valuable test grounds for sociolinguistic studies. In this section, I review the major discussions and related koinéization studies in the Chinese contexts against the background of rapid urbanization and industrialization. Some of the studies will be discussed in more detail because they are not published in English.

The notion of sociolinguistics was not introduced to China until the early 1980s when some dialectologists first started to pay attention to sociolinguistic issues in the newly arising industrial cities. Liang (1985), for example, investigated the linguistic situation of Dukou city (now Panzhihua city) in Sichuan province, where immigrants from mostly Dongbei (dōngběi 东北 ‘northeast’, usually refers to the three provinces Heilongjiang, Liaoning and Jilin in northeast China), Shandong, Hebei and other provinces arrived in the 1960s for the development of this steel industrial city. Liang (1985) reported different dialect zones found in Dukou, where the dominant dialect groups maintained their dialect in a closed community, usually a work unit (dānwèi, 单位), like a company or a factory (introduced below). However, when it came to the second generation, no matter what dialects their parents spoke or which dialect group they grew up in, they all used “Sichuan Pǔtōnghuà” as a lingua franca which combined features of their groups’ native dialects and the local

Sichuan dialect. Liang's work (1985) was considered to be the first paper that looked at sociolinguistic issues in new industrial cities in China (Lan, Cai & Zuo, 2010). Although no explicit references to the concepts were made, the situation of Dukou as described by Liang, from the current sociolinguistic point of view, was a typical koinéization process, where a new koiné "Sichuan Pǔtōnghuà" was formed through contact between different Mandarin varieties. However, at this stage, studies like Liang's (1985) were still conducted under the traditional framework of Chinese dialectology, in which data collection and sampling methods were seldom introduced, and analysis was based on only a few informants' data, which caused problems when the methodology in dialectology was employed in sociolinguistic research (Zhou, 2010, p. 70).

In the 1990s, another Chinese scholar, Yang Jinyi, began to promote sociolinguistic research in industrial cities on basis of his work in Luoyang, the capital city of Henan province. Luoyang began to develop as an industrial city from the 1950s, when immigrants from different dialect areas flooded in. Three residential districts, the old town district, the industrial district and the mixed district, were formed in the urban area due to different demographic situations. Residents of the old town district were mostly local people speaking the local dialect. The industrial district was filled with newly-built state-owned factories where most of the employees were immigrants. The mixed district was located between the other two districts, and was influenced by cultural and linguistic situations from both sides. Yang's study of Luoyang was conducted in the early 1990s. He investigated language use situations of the three districts in a number of different approaches, including direct methods such as questionnaires and interviews, as well as such indirect methods as observation of behavior in conversations. Apart from these, he was also able to take advantage of the household registration data from the police station to collect demographic information of more than 90% residents from the three districts (Yang, 1997), which was very rarely seen in sociolinguistic studies. Yang (1997) found that the linguistic situations of the three districts were very different. The old town district residents maintained their local dialect, whereas in the industrial district, a new variety of Pǔtōnghuà, Luoyang

Pǔtōnghuà, was formed in the second and third generations of the immigrant community. In the mixed district, the local dialect and Pǔtōnghuà varieties coexisted with a large proportion of bi-dialectal speakers, and the linguistic situation was rapidly becoming closer to that of the industrial district from the 1960s to 1990s. The process of how Luoyang Pǔtōnghuà was formed by dialect contact was similar to the formation of the Sichuan Pǔtōnghuà in Dukou, and could be taken as a koinésation process as well. However, as the main focus of his work was more towards language choice rather than specific linguistic features, the linguistic variables of the new variety were not examined in a detailed way. However, Yang (1997) reported some overall descriptive features of the Luoyang Pǔtōnghuà by presenting some examples: it was phonologically closer to Pǔtōnghuà, but adopted a number of frequently used words from the local Luoyang dialect, and the adopted words were often converted to Pǔtōnghuà tones (p. 60-61). These features are very similar to Hū Pǔ in the present study, in which the local lexicon l-words are adopted by the migrants, but some segmental features of these words are converted to the Pǔtōnghuà counterparts (see sections 3.1 and 4.4.1 for further discussion).

Based on these results, Yang later expanded this study by conducting larger-scale investigation among different work units in Luoyang and other industrial cities in neighboring areas in central China. Similarly to Liang (1985), Yang (2002) also found different work units to be closed speech communities. According to his data, in a work unit speech community, whether the second generation used Luoyang Pǔtōnghuà or the local dialect was closely related to the proportion of cadres and technicians among all employees. Yang proposed that this was because the cadres and technicians were mostly immigrants from more developed areas with higher education and social status. Their choice of speaking Pǔtōnghuà as a lingua franca in the work unit would largely influence the linguistic situation of the speech community.

Yang's studies can be considered as a very important step for the sociolinguistic research of Chinese industrial cities. First, he emphasized the sociolinguistic value of investigating immigrant cities developed under the background of rapid industrialization and

urbanization in China. Instead of only focusing on the language situation of a single city, Yang's projects attempted to find common patterns of linguistic change in Chinese industrial cities overall and relate the patterns to the social structure (see his work discussing overall language situations in Chinese industrial cities, e.g., Yang, 1999; 2002; 2004). Second, he advocated for the importance of work units as separate speech communities. I have mentioned the concept of the work unit or *dānwèi* (单位) a few times in this thesis so far. The social organization in China, very different from western societies, was based on the work unit structure, which "linked one's home with place of employment" (Jankowiak, 2013, p. 57). A work unit could be "a factory, shop, school, hospital, research institute, cultural troupe, or Party organ etc" (Lu, 1998, p. 53), which provided residency, labor insurance, health benefits and many other resources as combined social, political and economic institutions (Walder, 1986; as cited in Jankowiak, 2009). Therefore, a work unit could be considered as a very close-knit social network, which was likely to form different speech communities in industrial cities. Yang's work was not the first to look at linguistic situations within work units, but he emphasized the role of the work unit as a special social organization in Chinese societies and tried to explore the social factors influencing the outcome of dialect contact situations in different work units by conducting a large-scale sociolinguistic survey. Another important improvement in Yang's work could be seen in the methodological aspect. Unlike the previous dialectological studies, the methodology adopted in his research was much closer to the western sociolinguistic framework in terms of the sampling method and the use of questionnaires, interviews and observation. The use of the household registration system also provided him with valuable data. Quantitative methods were also attempted, but only very basic statistical analysis was conducted, by calculating basic number counts and percentages.

After the year 2003, sociolinguistic studies in China stepped into a new stage after the founding of the Sociolinguistic Society of China and the publication of *The Journal of Chinese Sociolinguistics* (Zhou, 2010). At that time, several leading western scholars such as William Labov, Gillian Sankoff and Anthony Kroch visited China and gave lectures on variationist

sociolinguistics. A number of Chinese scholars who received their PhD education in western universities also started to promote sociolinguistic research in China. With these efforts, the theories and methodologies from the western variationist sociolinguistic framework were introduced into the Chinese contexts. At the same time, Xu Daming established a sociolinguistic laboratory at Nanjing University, where a number of large-scale variationist sociolinguistic projects were conducted and a group of scholars was trained. The laboratory also held annual workshops on the topic of the “Urban Language Survey”, which provided opportunities for international exchange between sociolinguists in China and overseas. As Xu (2006) put forward, the area of Urban Language Survey (hereafter, ULS) integrated methodologies from dialectology, ethnomethodology, sociology, and especially quantitative sociolinguistics. Thereafter, studies on language/dialect contact, language change and variation in urbanization and industrialization in China were boosted. Most of the works were published in *the Journal of Chinese Sociolinguistics*, and another collection of recent publications was from the edited book *Industrialization and the Re-structuring of Speech Communities in China and Europe* (Van den Berg & Xu, 2010), and a special issue on Restructuring Chinese Speech Communities in the *Journal of Asia Pacific Communication* (Van den Berg, 2016).

From this background, a number of industrial and immigrant cities have been studied under the theories of koinéization since the beginning of the new century. One of the main projects from Xu’s team in Nanjing University was the study of the Kundulun District of Baotou. Baotou is a neighboring city of Hohhot, which is the largest industrial city of Inner Mongolia. The Kundulun District began to develop in the 1950s after it was designated as the residential area of the Baotou Steel Corporate, which was one of the biggest state-owned companies at the time. Immigrants in this district were mostly from northern China, especially the Dongbei (northeast) area, who formed their own closed community which resulted in contact-induced linguistic change. Xu’s PhD thesis (1992) was based on his first fieldwork in Kundulun in 1987. After the sociolinguistic laboratory was built, he took groups of investigators to Baotou for data collection again in 2003, 2004 and 2006, and built a real-

time database of recordings from more than 200 speakers. They investigated three phonological variables related to the nasal rhymes (such as /an/²), and quantitatively explored linguistic and social constraints operating on these variables (see e.g., Xu Daming, 1992, 2010; Xu Xiaohui, 2008; Wu, 2006; Liu, 2009). The results showed that the nasal rhyme variations were affected by such linguistic environments as onset, vowel, stress, tone; and also by social factors such as speakers' socio-economic status, original hometown, social networks, and occupation. Real time analysis found that compared to the 1987 data, both linguistic and social constraints for the variables had increased in number, which indicated increased uniformity in the behaviors of the members of the speech community (Zhu, 2010; Liu & Xu, 2012; Xu, 2010). For example, the nasalization variation was found to be predicted by occupation, social networks and place of origin in 1987, but the 2006 data showed that apart from these three factors, speech style, gender, age and education all became significant predictors of this variable (Xu, 2010). The effects of most linguistic factors were still consistent with the 1987 data, but some of the social constraints had changed. The factor place of origin was less relevant, but occupation became a stronger predictor (Wu, 2006). Liu and Xu (2012) also found that the locally-born generations were leading the change in the nasal final variations, and the shrinking influence of the place of origin factor implies the formation of local vernacular features in the Kundulun area. Another investigator Xu Xiaohui in Xu's team attempted to apply koinéization theories to the Kundulun context (2008). He claimed that a new local vernacular had formed induced by the dialect contact situation, which was named by local people as the Kundulun dialect (Kūn Qū Huà, 昆区话). By examining dialect choice behavior in Kundulun, he found that Pǔtōnghuà and the new variety of Kundulun dialect were the main chosen dialects, and the

² The three variables examined are 1) nasalization of the vowel preceding the nasal: [an] realized as [ã̃n]; 2) Omission of the nasal consonant: [an] realized as [a]; 3) Retroflexion of the vowel (or the r-colored vowel): [an] realized as [ar].

main factors that notably influenced the choice were age group/generation and occupation, directing to a typical and apparent social stratification.

The project of the Kundulun District could be seen as the first micro sociolinguistic study of Chinese industrial cities that looked at language variation and change using quantitative methods. However, the social constraints examined were mostly demographic factors. Xu Daming (2004), when redefining the notion of a speech community, emphasized the role of the speakers' identity in the formation of a speech community, and claimed that empirical and quantitative studies of attitudes should be done to further prove its significance (p. 21).

Another project from the Nanjing University research lab that did look at the role of attitudes in koinéization was the Jiangnan oilfield project, which was started in 2007 by Sun Deping (e.g., Sun, 2011; 2013). The Jiangnan oilfield was an industrial area in Hubei province, which was built in the 1960s and 1970s with a large number of skilled immigrants from all parts of the country. Similarly, a new local vernacular was formed as the lingua franca in this community, known as the oilfield dialect (Yóu Tián Huà, 油田话). According to Sun (2013), most oilfield residents believed that they were speaking Pǔtōnghuà with some local linguistic features. For example, they did not distinguish between /n/ and /ŋ/ in nasal rhymes, /n/ and /l/, /s/ and /ʃ/ at initial positions. By examining phonological features of the oilfield dialect, Sun (2013) found that it was gradually moving towards the standard Pǔtōnghuà after three generations of dialect mixing and leveling.

Next to the oilfield was the city of Qianjiang (潜江) where most residents spoke the local Qianjiang dialect. However, as Sun (2011) reported, the oilfield immigrants did not converge to the Qianjiang dialect although they were in fact surrounded by it, forming their own “dialect island”. He attributed this linguistic situation to two reasons: 1) The social interaction between the oilfield and Qianjiang residents was very limited. The oilfield was a highly self-sufficient community where people's needs could be easily met without leaving the area. The tall walls and iron gates around the oilfield also physically prevented most outsiders from entering the area. 2) The oilfield immigrants had much higher socio-

economic status than the Qianjiang local residents, and held strong prejudice against the Qianjiang dialect and communities. An interesting example given by Sun (2009) was the pronunciation of the word “Qian”(潜) in the place name “Qianjiang (潜江)”. “Qian” had tonal variation between the second tone “35” and the third tone “214” in Pǔtōnghuà, but the second tone was stigmatized because it sounded similar to the local Qianjiang dialect. The oilfield speakers were more likely to use the third tone in careful speech to avoid being recognized as Qianjiang dialect speakers. This is similar to the case of Høyang mentioned in section 2.2, where speakers’ identity construction behavior is presented in their pronunciation of the local place name. In this social context, Sun (2011, 2013) explored 476 oilfield speakers’ attitudes and their language use/choice using questionnaires in his fieldwork. He measured speakers’ attitudes towards Pǔtōnghuà and the Qianjiang dialect respectively using Likert scales of four dimensions: 1) pleasantness; 2) closeness and intimacy; 3) usefulness; 4) status and prestige. A value from 1-3 was then given according to each response, and the average score of the four dimensions was calculated as the final attitudinal score for each speaker. Sun (2011) found a significant correlation between speakers’ attitudinal score and their language choice. Speakers who demonstrated more positive attitudes towards Pǔtōnghuà and more negative attitudes towards the Qianjiang dialect tended to use Pǔtōnghuà more frequently in their daily lives. He thus argued for the role of speaker attitude in language change in the oilfield community. However, Sun did not explore the effects of attitudes on specific linguistic features, and the measurement of attitudes was also rather simplistic. Moreover, his questionnaires only focused on attitudes towards the two varieties and lacked an in-depth investigation of speakers’ identities, social networks, and other aspects of their attitudes. Despite these, the Jiangnan oilfield project was still valuable in its attempt to explore socio-psychological factors in koinéization of the Chinese context.

These projects conducted by the Nanjing University sociolinguistic lab could be regarded as pioneer studies that applied theories and methodologies from variationist sociolinguistics to koinéization situations in China. Quantitative methods and statistical modeling were

used to explore the effects of social and linguistic factors on language use and specific linguistic variables. However, the main statistical tools used in these studies were the Variable Rule Program (VARBRUL) and GoldVarb (e.g., Cedergren & Sankoff, 1974; Rand & Sankoff, 1990), which received more criticism in recent years especially after the use of mixed effects regression models was popularized. For example, VARBRUL and GoldVarb could not model continuous factor groups like age, and the procedures exploring interaction effects were too complicated. Moreover, these tools could only model fixed effects like speaker sex, age, social class, but were not able to take potential speaker-level and word-level variation into account (Johnson, 2009; Tagliamonte, 2011; see a more detailed discussion of this in section 4.8).

Apart from works of the Nanjing research team, another set of recent literature studying koinéization in China used descriptive approaches, for example, the Qingshan District study in Wuhan (Yang, 2010; Lu, 2014), the Shangrao “railway dialect” study in Jiangxi (Yang, 2013); the Pucheng “Ou dialect” island study in Zhejiang (Chen, 2013); the Changanying island study in Hunan (Tang & Li, 2016), among others. Most of these studies were thesis work from Fudan University, supervised by Prof. You Rujie and Prof. Tao Huan. The methodologies of these papers were still mainly similar to dialectological work, so they did not use quantitative methods to explore sociolinguistic variation. However, they provided very detailed description of the linguistic and social situations of the localities, and examined the inter-generational language changes in different stages of the koinéization process. The new varieties/koinés formed in these localities were very interesting and should definitely be investigated more in variationist sociolinguistic approaches in future studies. For example, Lu (2014) reported the “Wān Guǎnzi accent” (弯管子话) formed in Qingshan District in Wuhan province, which was built in 1952 as an industrial district. “Wān Guǎnzi” in Pǔtōnghuà literally means “a bent tube”, but the local Wuhan people used this term to refer to the mixed variety spoken by immigrants who tried to adopt the local accents but failed. Lu (2014) explained that the word “wān” (弯, bent) here means weird or awkward, while “guǎnzi” (管子, tube) refers to the speech organs like the throat. Lu

investigated four generations of the Qingshan residents, and found that this mixed variety began to form in the second generation and became more focused and stable in the third and fourth generations. However, the fourth generation was using this variety less frequently, choosing to speak Pǔtōnghuà or the Wuhan dialect on most occasions. Another interesting study conducted by Yang Wenbo (2013) was about the Shangrao railway dialect (上饶铁路话). This speech community was formed around a railway station in Shangrao, Jiangxi province, and all the first-generation residents were immigrants who also worked in this railway station. Yang (2013) found that the koinéization process in this community was completed in only two generations. This was because the majority (over 80%) of the immigrants were originally from the same province, Zhejiang, speaking similar dialects. They also formed a very isolated and closed community because of the railway system and the geographic location. The second-generation speakers all went to the same primary school and high school operated by the railway authority where only children from the railway community could be enrolled. In this social background, the railway dialect was formed at a very rapid speed in the second generation. However, according to Yang (2013), this mixed variety had become an endangered dialect now as the third generation had mostly shifted to speaking Pǔtōnghuà or Shangrao dialect. He argued that this could be related to the decreasing socio-economic status of the railway employees. Studies of koinéization processes in China like these have been increasing in recent years and have introduced various interesting social and linguistic contexts in Chinese societies. The role of socio-psychological factors like attitudes and identity was mentioned in some studies but was seldom explored in quantitative analysis.

At the same time, along with the popularization of variationist sociolinguistics in China, studies on language attitudes have also been promoted in recent years. However, instead of exploring the effect of attitude on specific linguistic variables, most of the works still focused on the relationship between language attitudes and language choice. For example, Guo (2007) interviewed 54 speakers in Lishui, a small town in Nanjing suburban area, and investigated their attitudes towards the local Lishui dialect, Pǔtōnghuà, and some other

varieties. He found that Pǔtōnghuà was afforded a very high prestige among the local residents. Meanwhile, he also provided numerous examples to show how the local speakers were trying to accommodate to Pǔtōnghuà in phonological, lexical and syntactical aspects, and claimed that this could be related to their positive attitudes towards Pǔtōnghuà. Another study conducted in Nanjing by Zhou (2011) explored 292 speakers' attitudes towards the Nanjing dialect and Pǔtōnghuà using questionnaires asking simple questions like "Does the Nanjing dialect sound pleasant?" and "Is Pǔtōnghuà useful in Nanjing?" She then investigated the effects of social factors on their language choice using simple statistics like numbers and percentages, and found that speakers' language choice was sometimes consistent with their language attitudes, but sometimes not, which was related to other social factors like speakers' sex, age, occupation and region. Tang (2016) was another large-scale language attitudes study in which more than 1800 questionnaires were investigated in Shenzhen, Guangzhou. Tang claimed that the main varieties spoken in this area: Cantonese and Pǔtōnghuà, both had their own strong support groups in terms of language attitudes. Cantonese was highly recognized by secondary and university students, while Pǔtōnghuà was more related to higher education and hi-tech jobs in work force communities. However, Tang did not look at the correlation between speakers' attitudes and their linguistic behavior. Another recent study by Jing and Zhu (2016) explored language attitudes and language choice in younger generation Beijing-born residents in Beijing. They interviewed 265 speakers about whether they would prefer the local Beijing dialect or Pǔtōnghuà and why. The results showed that more speakers reported they preferred the Beijing dialect than Pǔtōnghuà, and their reasons were mostly related to feelings of identity such as a sense of belonging. Reasons given by participants who preferred Pǔtōnghuà were more objective, usually related to its function as the standard language. Jing and Zhu also investigated the social constraints operating on these speakers' linguistic behavior (accent-switching, self-reported and observed accent choice) using statistics, and found a significant effect of speakers' social networks, which was measured using a list of the five most frequent contacts provided by each speaker (Jing & Zhu, 2016, p. 36). However, as speakers' attitudes were not quantitatively measured, they did not

examine the effects of attitudes, or the interaction of attitudes and social networks in the statistical models.

To summarize, research into language change caused by migration, urbanization and industrialization in Chinese urban cities has been developing very quickly especially since the beginning of the new century. Variationist sociolinguistic approaches have been taken to examine the social and linguistic constraints operating on language variation and language use, but studies using newly developed quantitative methods and statistical tools should be further promoted. The role of socio-psychological factors such as speaker attitude in this sort of language change has been discussed to some extent, but certainly needs to be further investigated in more empirical studies.

3 The locality, Hohhot

This chapter introduces the social and linguistic background of the locality of Hohhot, and explains how it is similar to or different from other localities studied in previous new dialect formation literature.

3.1 Hohhot and the emergence of Hū Pǔ

Hohhot (Hūhéhàotè 呼和浩特), with its name meaning “blue city” in Mongolian, is the provincial capital of Inner Mongolia Autonomous Region (IMAR) located in north central China. It is a second-tier city with a population of almost two million in the central urban area (2010 census, 2013). About 85% of the population are ethnic Hàn 汉 Chinese, and most non-Hàn residents of Hohhot speak Jìn dialect (Jìn Yǔ 晋语) or Pǔtōnghuà (Puthuval & Wang, 2017). Only a very small proportion of Mongols speak Mongolian and they are mostly bilinguals as well. This dissertation only focuses on the Chinese-speaking communities in Hohhot.

Officially, the urban center of Hohhot is divided into four administrative districts, but residents in Hohhot often regard the city as bifurcated, composed of the New Town and the Old Town (Jankowiak, 1993; see Figure 3.1).

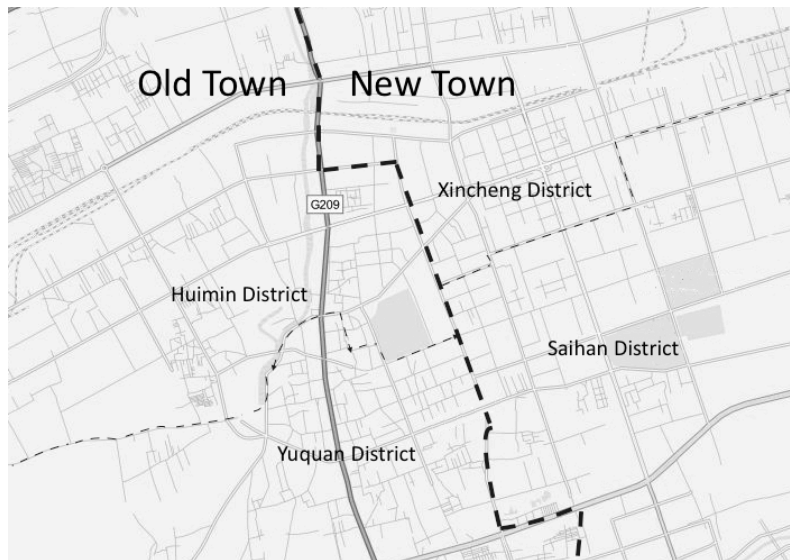


Figure 3.1 Map of the Hohhot urban area, with the New Town and the Old Town separated by the thick dashed line.

Like many other Chinese immigrant cities mentioned in section 2.4, Hohhot also began to develop in the background of China's industrialization and urbanization in the 1950s. Before 1950, the overwhelming majority of Hohhotians were mainly locally-born Hàn Chinese whose ancestors were peasants and migrated to Hohhot from neighboring Shānxī, Shǎnxī, and Héběi provinces (see Figure 3.2). Linguistically, they spoke a sub-variety of Jìn dialect. At that time, most local residents were living in the Old Town of Hohhot, and the New Town area was sparsely populated. However, after 1954, when Hohhot was designated as the capital city of Inner Mongolia, a large wave of migration began (Jankowiak, 1993, p. 17-19). At the same time, a great number of intellectuals, cadres and technicians from more developed areas arrived encouraged by the government policy to increase the population and the level of development in the frontier. As a result, it is estimated that, between 1950 and 1970, about 1.4 million people moved to urban Hohhot, nearly tripling the population (Hohhot, 1984). Most of the migrants were concentrated in the New Town area, forming their own communities. Originally they were from all parts of the country, speaking different regional dialects, but they used Pǔtōnghuà (with accents from their original hometowns) as a lingua franca in Hohhot.

Jìn dialect is different from Pǔtōnghuà on various linguistic levels. The features of Jìn dialect have been discussed at length in the previous literature, including phonological features such as the keeping of the entering tone³ (rùshēng 入声) marked by the glottal stop coda [ʔ], merging of tones, monophthongization of rhymes, merging of /tɕ tɕʰ s/ with /ts tsʰ s/ (descriptions of these features are available in English in e.g., Shen, 2017; for discussions in Chinese, see e.g., Liu & Liu, 2000; Hou, 1999); lexical features such as personal pronouns like /ŋɤ³¹/ 1SG ‘I’ and /nia³¹/ 3PL ‘other people’ (Lan, 2012), interrogative pronouns like /səŋ⁴³⁵/ ‘what’ and /tsa⁵³/ ‘how’ (Xing, 1960), kə-type prefix words (a set of disyllabic words with /kəʔ⁴³/ as the first syllable, see Shen, 2017), l-words (Hou, 1999; Shen, 2017); and syntactic features such as the frequent use of sentence-final modal particles like wā 哇 (Jin, 2008; Li & Xin, 1987; Xing, 1995), and a tendency to postpose adverbs like dōu 都 ‘all’, cái 才 ‘only, just now’, and fǎnzhèng 反正 ‘anyway’ (Li & Xin, 1987).

Generally, the new variety Hū Pǔ is phonologically closer to Pǔtōnghuà, while absorbing numerous lexical and grammatical elements from Jìn dialect. Therefore, the lexical features mentioned above such as the personal pronouns, kə-type prefix words, and l-words are often adopted by Hū Pǔ speakers. However, when they pronounce these words, Hū Pǔ speakers tend to convert the vowels and/or tones in these lexemes to the Pǔtōnghuà forms. For example, Hū Pǔ speakers pronounce the personal pronoun /nia³¹/ ‘other people’ as /nia³⁵/, converting the Jìn tone “31” to the Pǔtōnghuà equivalence “35”. This linguistic phenomenon is very similar to the Luoyang Pǔtōnghuà (Yang, 1997) discussed in section 2.4, in which a number of frequently used lexemes were adopted from the local Luoyang dialect, but the adopted words were converted to Pǔtōnghuà tones (see further discussion in

³ The entering tone, or the checked tone, is a distinctive tone category in ancient Chinese. Syllables in this tone category are short in duration with distinctive pitch contour, and contain stop endings – p, –t, –k, or the glottal stop ending –ʔ (Yun, 2017). The entering tone has merged with other tones in modern Pǔtōnghuà, but it is kept in many Chinese dialects, such as Cantonese and Min (閩) dialect. In Jìn dialect, the entering tone is marked by the glottal stop –ʔ.

section 4.4.1 and section 7.3). Also, Hū Pǔ speakers often show inter-speaker variations in their degree of adopting these Jìn lexemes. For example, some speakers use /nia³⁵/ as a third person pronoun very frequently while others might rarely adopt it but use the Pǔtōnghuà equivalent /zən³⁵tɕiɑ⁵⁵/(rénjiā, 人家) instead.

Therefore, in the new variety Hū Pǔ, these adopted lexemes are likely to contain numerous linguistic variabilities, which may lead to interesting sociolinguistic findings. The present study will focus on one set of these lexemes: l-words, which are used by speakers of different varieties in Hohhot, displaying variation in different linguistic levels (see section 4.4.1).

3.2 Sociodemographic background

The formation of Hū Pǔ can be considered as a process of koinéization. It is similar to the new town scenarios of new dialect formation (Kerswill & Trudgill, 2005), like the koiné found in Milton Keynes or Høyanger, because there were prior Jìn speakers in the area before the migration and intense social communication has been found between the Old Town and New Town residents. Siegel (1985) defines this sort of new dialect as an “immigrant koiné”, which has resulted from a mass migration of speakers and become “the primary language of the immigrant community” (p. 364). This section provides a wider context of the demography of Hohhot.

3.2.1 Geographic origins

The first-generation migrants in Hohhot were originally from all parts of the country with no obvious dominating dialect groups. As reported by Jankowiak (1993), “tens of thousands of Chinese from Nanjing, Shanghai, Tianjin and Beijing either volunteered to move to Hohhot, or were transferred, individually or collectively, along with their work units, to Hohhot” (p. 17). These cities he mentioned were major cities in China, where most of the

well-known universities were located. So many intellectuals who migrated to Hohhot from these cities were graduates from the local universities, but originally they could be from any part of China. The participants in my fieldwork also demonstrated this feature: the 13 first-generation migrants were originally from 10 different provinces throughout China (see section 4.2 and Figure 4.1). Therefore, the migrant community in the New Town of Hohhot was a very heterogeneous one at that time. This situation is different from the railway dialect formation in Shangrao, where more than 80% of the migrants were from Zhejiang province; or the koinéization process in Kundulun District of Baotou, where the majority of migrants were from the Dongbei area. In this sort of heterogeneous context, as mentioned in previous discussions of the Tyssedal koiné in Norway, the more standard linguistic forms are often more likely to survive.

3.2.2 Socio-economic status

The state-sponsored migrants in Hohhot were mostly well-educated cadres, intellectuals or skilled workers, as noted above. So the majority of them, after moving to Hohhot, worked as government officials, administrators, professors, schoolteachers, doctors or workers in state-owned factories. In contrast, the locally-born residents had a lower social status and were mainly handicraft workers, shop assistants, street peddlers, or workers from collective-owned factories (Borchigud, 1996). Therefore, there was a clear social stratification between the New Town and the Old Town. This situation is actually prevalent in most of the Chinese immigrant cities that arose in the 1950s during industrialization, such as the case of the Jiangnan oilfield, Luoyang, and Kundulun, because the state-sponsored migrants were supposed to be sent to the less-developed areas to help with the local construction.

3.3.3 Social networks, the work unit social structure

As introduced in section 2.4, the social organization in China is based on the work unit structure. The residential situation in Hohhot also followed this type of social structure,

where the employees from the same work unit were usually residents of the same neighborhood. This could be considered as a type of close-knit network. All the colleagues were also neighbors, who would visit each other, look after children for each other, and sometimes exchange food (Jankowiak, 1993). Their children grew up together, and most of them went to the same kindergarten then the same primary school; thus they naturally formed a neighborhood-based peer group, which was of crucial importance for the formation of the new dialect. This social context is different from some previously mentioned European localities, where communities were very isolated in sparsely populated areas, and children did not have a peer group to conform to, like the English Fens and the Falkland Islands.

Many informants in my fieldwork also regarded this work-unit neighborhood concept as a sort of label for themselves, claiming that they were from, for instance, “Gōngdà” (工大, abbreviation of the Inner Mongolian University of Technology), “Nóngdà” (农大, abbreviation of the Inner Mongolian Agriculture University), “253 Hospital”, or “Dǎngwěi” (党委, Party Committee). When they said that, it did not necessarily mean that they worked there, but it could be that they simply lived in that neighborhood because they were relatives of the employees.

According to Jankowiak’s (1993) observation, many neighborhoods were socially homogeneous, such as intellectuals’ work units in the New Town, and working class communities in the Old Town; however, there were also neighborhoods “owned and regulated by the municipal housing authority” (p. 110), where the two classes were mixed, and dialect contact could happen more frequently. Moreover, even within many homogeneous migrant communities, there were also interactions between different classes. For instance, all work units would have logistics departments which usually provided facilities like dining halls, public bathhouses, boiler rooms (guōlúfáng, 锅炉房) etc. This department recruited a number of blue collar workers and, according to many informants in my interviews, most of those blue collar workers were locally-born Jīn speakers, who were also likely to be their neighbors. Although they might be living in different

subsections of the neighborhood, they had opportunities to frequently interact in various situations, and their children would still form into the same peer groups. As Jankowiak (1993) has pointed out, children often “serve to facilitate easy and frequent interactions between individuals from different households” (p. 109). Therefore, although migrants from homogeneous communities and mixed-class communities had different degrees of integration to the local Jìn-speaking communities, the dialect contact situation was prevalent. Some first-generation migrants in my fieldwork also reported that they had experience of working in neighboring Jìn-speaking areas for a year or two, and some even married Jìn speakers, which gave them more opportunities to be exposed to the local dialect and culture.

From this aspect, compared to other Chinese immigrant cities, the social context of Hohhot is closer to that in Dukou or Luoyang, where the migrants and the locals are integrated to a higher degree, while the situation of Jiangnan oilfield or Kundulun was quite different, where the migrants were concentrated in their own closed communities. Therefore, it could be expected that the outcome of dialect contact, or Hū Pǔ, might be influenced more by the local dialect than the Jiangnan oilfield koiné or the Kundulun koiné.

3.3.4 Isolation from original hometowns

The migrant community in Hohhot was very isolated from their original hometowns after they moved there. This social characteristic is not unique to Hohhot, but is applicable to most Chinese industrial cities arising in the 1950s and 1960s. At that time, when telephones were not common for family use, the ways for communicating with distant families were mainly by mail or telegraph. In addition, most state-sponsored migrants were in their 20s or 30s, just graduating from college or at the early stage of their career. Their financial status did not allow them to visit their places of origin frequently, especially those whose hometown was thousands of kilometers away in the very southern part of China. One informant in my fieldwork reported how her family strived to save up money to visit her hometown in Dongbei (northeast) once every six years, which, according to her, was very

frequent compared to other migrants she knew, who could have no chances to go back at all. For the second generation migrants, most of them had never been to their parents' hometowns. This degree of isolation is completely different from what Mæhlum (1992) had found in the Spitsbergen koinéization, as mentioned in section 2.2, where the migrant families went back to their place of origins for summer breaks. The Spitsbergen second-generation children were found to be strongly influenced by their parents' speech, but the situation of Hohhot, as we will see later, is quite different.

3.3 Social conflicts between the New Town and the Old Town

3.3.1 Evidence from anthropological literature

The mass sudden influx of migrants in Hohhot, and the clear distinction of socio-economic status between the two towns, imposed many social issues and conflicts between the locally-born Jìn-speaking residents and the newcomers (Borchigud, 1996, p. 164). Jankowiak (1993) pointed out the new urban migrants' "ingrained pity at the plight of the Old City's (Old Town's) residents, who were repeatedly referred to as 'backward, dirty, and feudalistic'" (p. 17). With the rapid development of the New Town area, Hohhotians came to associate the New Town with "modernity, material success, and the good life", while the Old Town was perceived, by contrast, as "an increasingly backward, worn-out, and pitiful place to live" (p. 10). In this social context, the local Jìn dialect, as spoken by the locally-born communities, was also stigmatized. Borchigud (1996) mentioned "many Mandarin-speaking Han migrants in Hohhot during the 1960s and 1970s discriminated against the local Han

dialect speakers ... by calling them *cǐdì kuǎzi* (此地侬子, the local-accent speakers⁴)” (p. 165). As reported by Jankowiak (1993), these sorts of conflicts and prejudice were reflected even more in those heterogeneous neighborhoods where the two classes interacted on a daily basis. Most of the intellectuals maintained “a polished veneer of respectability (i.e. gentility) in their daily interaction with working-class neighbors”, which “disguises a genuine contempt” for them, and they were reluctant to develop any kind of warm, close relationship with them (p. 111). However, this was not the case for everyone in the migrant community. Jankowiak (1993) also mentioned that there were a few intellectuals and cadres who resented the “fences” between classes, and showed “willingness to interact warmly with their working-class neighbors” (p. 112).

This evidence from previous anthropological works indicates that the social issues between the two towns and the prejudice towards the locally-born communities have been long-standing and could have emerged in the first-generation migrants. However, the social situation described in these works was Hohhot in the 1980s and early 1990s, which might have changed in the following decades.

3.3.2 Evidence from local online forums

In order to better understand Hohhotians’ social-psychological orientations today, I collected some attitudinal information from the online forums before the fieldwork. I looked for posts about Hohhot people’s opinions of the local community on some popular online forums in China such as Bǎidù Tiēbā (百度贴吧) and Tiānyá (天涯), where people can post their opinions about a certain topic and provoke open discussions. About 50 topics were found concerning Hohhot people’s attitudes towards Jìn dialect or Old Town communities, and many of the posts still revealed the social conflicts between the migrants

⁴ This translation is from Borchigud (1996). 侬子 *kuǎzi* refers to speakers who have strong accents, and it usually has a derogatory connotation.

and the locals. Some of the posts showing typical prejudice against locally-born people are presented below in (3-1).

(3-1)

“You have to admit that Old Town people are picked at because they are bourgeois, uncultured and always keen to gain small advantages” (Xiaoduantuixiaoyanjing, 2014).

“I’m in frequent contact with Old Town people. Many of them are calculating, not open enough, and another defect is snobbery” (Jingxingjiahe, 2011).

“I was born and brought up in the New Town area. I can understand Jìn dialect, but will never speak it. It sounds interesting and hilarious, but it’s unpleasant to me” (Jinchajifangyan, 2011).

“Jìn speakers are either from suburban areas or the less-developed Old Town. If children speak Jìn dialect at school, they’ll be laughed at” (Didaojiuchengren, 2014).

Some posts indicate that this prejudice is even influencing local people’s choice of marriage partner. As is shown in the post in (3-2), Yanyanaichitang explicitly claimed that she would not marry a Jìn speaker.

(3-2)

“I won’t say I hate Jìn dialect, but I’ll never marry a Jìn speaker. Is there anything wrong with choosing someone you like in marriage?” (Yanyanaichitang, 2013)

However, not everyone demonstrated this sort of prejudice. Some posts by New Town people defended it for the locally-born residents in the discussion about the “stigma” of the Old Town, and emphasized the traditional and cultural value of the Old Town, as shown in the examples in (3-3).

(3-3)

“Old Town people are not bad. They are down-to-earth, dutiful, realistic, and sometimes come forward for friends. They represent the traditional Hohhot culture” (Laowannao09, 2011).

“Although I’m not from the Old Town, I have many friends from there. Old Town people are the indigenous residents in Hohhot. No Old Town, no Hohhot!” (Buchizhuroudecaoyuanlang, 2010)

“In fact, the earliest Hohhot culture lies in the Old Town, which is the purest”(Yuchenshang, 2014).

“For the Old Town, it’s a shame that many traditional houses and buildings were destroyed in the modern construction. Those old houses can actually represent the lifestyle of Hohhot before the 1950s. If they are properly protected and renovated, they could make a great cultural area, which will be amazing. It’s really a shame” (Yayayaya, 2010).

The posts collected from the online forums were informative in revealing the current socio-psychological orientations of the Hohhotians. Therefore, they were later used as one of the data sources for the design of the attitudinal questionnaires (see section 4.5.1). However, the limit of this data source is also obvious, because people who posted these opinions are most likely to be young people who are familiar with computers and the Internet. The type of forums investigated might also be restricted to a select group of people.

Despite that, it can be seen from these posts that, even today, the social conflicts between the locally-born and the migrant communities are still widely found. This prejudice against the Old Town has been deeply rooted in some of the New Town residents’ minds. However, there were also people from the migrant community who presented neutral or positive attitudes towards the local dialect and culture. Therefore, Hohhot is an interesting locality for sociolinguistic studies concerning attitudes.

4 Research design and methodology

The data collection was conducted in Hohhot from August to October, 2014. Speakers representing three generations from both the New Town and the Old Town were interviewed, with their linguistic, attitudinal and social networks data collected. This chapter introduces the field methods and elaborates on the linguistic and social variables.

4.1 Entering the community

I was born and grew up in the New Town area of Hohhot, so entering the community was not too difficult. As an indigenous Hohhotian, I could easily take the position of an “insider”, and locating informants started quickly even at the beginning of the fieldwork. However, the participants were very unevenly distributed at first due to the limitation of my own social networks and background. Most of the early-stage informants found were younger and middle-aged speakers from the New Town who were introduced by my friends and schoolmates, as well as my parents’ friends and colleagues. In order to reach the communities beyond my own social circle, I also frequently went to various social events during my fieldwork trip, like weddings and school reunions, which allowed me to connect to a wider social network, and was proved to be an efficient way to find more participants. The recruitment of older participants was more of an issue than I had expected. The first-

generation migrants were mostly over 80 years old at the time of the fieldwork, so many potential participants in my own social circle had passed away or were unable to talk too much in the interview because of their health condition. This issue was later solved with the help of my aunt. She was a popular and warm-hearted person in her neighborhood or work unit: the Gōngdà university (工大 Inner Mongolia University of Technology), where most of the retired professors were state-sponsored migrants originally from all parts of China. As a member of this retired professor's community, my aunt kept up very frequent communication with these colleagues and neighbors by their visiting and helping each other on a daily basis. Many of them also had a routine of getting together in the residential streets after dinner to play cards or chess, or simply chat. Therefore, with my aunt's help, I successfully interviewed many state-sponsored migrants within a short time. The importance of finding the "key person" in the community has often been emphasized in ethnographic fieldwork (e.g., O'Reilly, 2005). My aunt acted as this key person in my fieldwork to enable me to gain access to those communities beyond my social circle. Moreover, since I was introduced as a relative of my aunt, all her friends were willing to spend hours talking about their life stories, which was crucial for the current study in terms of eliciting more genuine attitudes.

Finding participants from the Old Town was difficult at the beginning as well because I had few personal contacts there, but later, once after I interviewed a few of them, the "snowball" strategy could be used and people were usually willing to introduce me to their friends or neighbors. However, it was still an issue for me to access the older generation. So at the later stage, as a secondary solution, I went back to my aunt's social networks, and found some older Jìn speakers who were born and grew up in the Old Town, but later moved to the New Town. This strategy helped with the collection of a more complete data set of the Old Town speakers, but we have to bear in mind that these older speakers cannot be considered as representative of the Old Town. Many of them were also retired professors or engineers from the Gōngdà community, whose attitudes and language might be different from the typical working-class residents of the Old Town.

4.2 Participants and data collection procedure

During the 10-week fieldwork, altogether 67 speakers were interviewed and recorded. 35 participants were from the New Town and 32 were from the Old Town⁵. Participants were evenly distributed between male and female, and across three age groups. For the New Town participants, the older speakers were first-generation state-sponsored migrants, who came to Hohhot in the 1950s or 1960s, and they were born between 1930 and 1941. The middle-aged group was the second-generation migrants who were born in Hohhot between 1949 and 1962. The younger group was the third generation born between 1985 and 1995. Most of the New Town participants were state-sponsored migrants and their offspring, whose original hometowns were reported to be Běijīng, Tiānjīn, Shāndōng, Liáoníng, Guǎngdōng, Jiāngsū, and other provinces (see Figure 4.1). The Old Town participants were roughly at the same age levels. The distribution of participants is shown in Table 4.1.

Table 4.1 Participants' distribution.

Town	New Town						Old Town					
Age	Older		Mid-aged		Younger		Older		Mid-aged		Younger	
Sex	M	F	M	F	M	F	M	F	M	F	M	F
No.	7	6	4	6	5	7	5	4	5	7	5	6
Total	67											

All participants were interviewed either in their own homes, or other quiet environments. At the time of the interview, all of the older informants and most of the middle-aged informants had retired, so they often preferred to be interviewed in their own homes. Many

⁵ Speakers who had experience of moving between the two towns in their lifetime were grouped according to their childhood experience. If they grew up in the Old Town, they were considered to be Old Town participants even though some of them were living in the New Town at the time of the interview.

times when I entered a participant's home, I was happy to find that his/her family members, or a visiting neighbor, were also willing to participate in my study, so this actually helped me to find more speakers. The younger participants usually chose to come to my home or a small classroom in a local university or some public areas like cafés. Some participants were interviewed individually and others were interviewed in groups of two or three people. Recordings typically lasted for 60-90 minutes and all procedures were recorded using a Sony ICD-TX50 digital voice recorder.



Figure 4.1 Original hometowns of the state-sponsored participants in this project, with migration movement directions indicated by the dotted arrows. (Map retrieved January 15, 2015 from https://en.wikipedia.org/wiki/File:China_administrative_claimed_included.svg)

The data collection phase comprised several parts. A word elicitation task was conducted at the very beginning to collect linguistic production data, when speakers did not have any idea about which linguistic feature I was examining, so that the production could be more

natural. After that, speakers' attitudinal data were collected using questionnaires, followed by a semi-structured interview, in which I asked the participants to talk about their opinions of the local communities, different dialects in Hohhot, and their own accents. The last part was to collect their social networks data. These parts of the data collection procedure are elaborated on below in sections 4.4, 4.5 and 4.6.

4.3 My own stance in the community

Since this study has a particular focus on attitudes and language, my own position in the local communities and my linguistic production have to be cautiously dealt with in the interviews.

As an indigenous Hohhotian who grew up in the New Town area, I could easily talk about attitudes towards the Old Town and Jìn dialect with my New Town participants. But when it came to the Old Town participants, it is possible that they might hide their true feelings and avoid talking about their negative attitudes towards the New Town because of my identity. However, my own family background allowed me to use an “insider” strategy. My paternal grandparents were state-sponsored migrants originally from Jiangsu province, whereas my maternal grandparents were Jìn speakers originally from Shanxi province. This “mixed” heritage allowed me to “play off” both sides. When interviewing New Town participants, I would emphasize my own experience in the New Town, and the state-sponsored- or migrant-related aspects of my background. By contrast, when interviewing Old Town participants, I usually mentioned my maternal ancestors first and emphasized that we were of the same origin, but I seldom talked about my New Town background.

In terms of language, in order to reduce the possible influence of my own accent to the minimum, I chose to speak a mild-accented Hū Pǔ in all the interviews and was trying to be

consistent⁶. Some participants even commented on my accent: some New Town speakers pointed out my “obvious” Hū Pǔ accent and admired my “good” imitation of the local Jìn dialect, whereas Jìn speakers usually laughed at my “bad” imitation.

4.4 Linguistic variables

As introduced in section 3.1, Hū Pǔ absorbed numerous lexemes from Jìn dialect. Among these adopted Jìn elements, this study particularly focuses on a special set of lexemes: l-words (Hou, 1999), because they display linguistic variabilities that are potentially linked to different varieties spoken in Hohhot, which allows me to detect to what degree individual speakers adopt Jìn features. This section introduces the linguistic feature l-words and the methodology for linguistic data collection.

4.4.1 L-words

L-words are a type of disyllabification, whereby a monosyllabic word splits into a disyllabic word (Shen, 2017), with /l/ between the onset and rhyme, hence the term “l-words”. The typical pattern of l-words, as shown in the examples in Table 4.2, is for the first syllable to have a reduced vowel and end in a glottal stop, and the second syllable to have /l/ as the initial sound. Shen (2017) formulated this process as $C_1VC_2 \rightarrow C_1\text{ə}ʔ-lVC_2$. Note that this is not an absolutely strict word-formation rule. There are also variabilities in the l-word lexicon. For instance, if the monosyllabic words have the sounds [i] or [u] immediately following the onset, the [i]/[u] is usually retained in the first syllable of the disyllabic form as /uəʔ⁴³/ or /iəʔ⁴³/, like in /xuəʔ⁴³la³⁵/ 划拉 ‘scribble’ and /t^hiəʔ⁴³liu³¹/ 提溜 ‘carry something randomly’.

⁶ This is also very close to my typical speech pattern when I am in Hohhot.

Semantically, l-words usually gain a more concrete and special meaning than the original monosyllabic words (Sun, 2006).

Table 4.2 Some examples of l-words in Jìn dialect and their corresponding monosyllabic forms (words selected from Ma & Xing, 1997).

Disyllabic form (l-words)	Monosyllabic form
/xuə ⁴³ la ³⁵ / 划拉 ‘scribble’	/xua ³⁵ / 画 ‘draw’
/pə ⁴³ lai ³¹ / 卜徠 ‘swing randomly’	/pai ³¹ / 摆 ‘swing’
/k ^h uə ⁴³ luŋ ³¹ / 窟窿 ‘holes’	/k ^h uŋ ³¹ / 孔 ‘holes’

L-words are considered to be a salient feature in all sub-varieties of Jìn dialect; Ma and Xing (1997) reported more than 70 l-words spoken in the Inner Mongolia area. Some l-words can only be found in Jìn dialect, while others are also used in spoken Pǔtōnghuà, and are pronounced differently on many linguistic levels, such as their vowels, tones, consonants, stress patterns.

As discussed in section 3.1, Hū Pǔ speakers, when adopting lexical features from Jìn dialect, often converted the salient segmental features like tones or vowels into their Pǔtōnghuà counterparts. This is also the case for the adoption of l-words. For example, the l-word 合浪 ‘an alley or a lane’ is pronounced as /xə⁴³lā³⁵/ in Jìn dialect, but Hū Pǔ speakers are very likely to produce it as /xə⁴³laŋ⁵¹/, in which the Jìn tone “35” is changed to the Pǔtōnghuà counterpart “51”, and the vowel /ā/ is produced in the Pǔtōnghuà form as /aŋ/.⁷ This may relate to the fact that Jìn dialect has very different vowel and tone systems compared to Pǔtōnghuà, so these features may be too salient to be adopted. However, in this study, in order to detect the speakers’ degree of adopting local features, it is important to select a linguistic feature that demonstrates inter-speaker variation.

⁷ The Pǔtōnghuà diphthong /aŋ/ is nasalized and monophthongized as /ā/ in Jìn dialect, which is a salient Jìn-feature (see Shen, 2017).

4.4.1.1 The stress pattern variable

Stress pattern is such a feature according to my observation and experience as a Hū Pǔ speaker myself. As mentioned in the introduction, l-words, as disyllabic forms, display variation between a weak-strong pattern and a strong-weak pattern. The weak-strong pattern is a typical Jìn feature, whereas the strong-weak pattern is associated with Pǔtōnghuà. For example, the l-word 划拉 ‘scribble’ is pronounced as /xuəʔ⁴³la³⁵/ in Jìn dialect with a reduced vowel and an entering tone in the first syllable and a full tone in the second, which is perceived as a weak-strong pattern. By contrast, the same l-word is pronounced as /xua³⁵la⁰/ in Pǔtōnghuà, with a full tone in the first syllable and a neutral tone in the second, displaying a strong-weak pattern (see a more detailed description of this feature in chapter 5).

Hū Pǔ speakers, who have been exposed to the pronunciation of l-words both in Jìn dialect and Pǔtōnghuà, use weak-strong and strong-weak stress patterns variably when they produce these l-words, thus forming a binary linguistic variable. Since the two types of stress patterns are associated with different language varieties (Jìn and Pǔtōnghuà), this variable is perfect for me to detect individuals’ adoption of Jìn features, and could be a potential testing ground for understanding the effects of attitudes on their linguistic production.

In this thesis, another issue brought up in research question 6 is related to the level of awareness or relative salience of the linguistic variables. That is, whether the effects of attitudes on language could be found if speakers are not explicitly aware of the linguistic variable. The stress pattern variable is not a good candidate in terms of this issue, because local speakers explicitly understand that the weak-strong pattern is a salient feature of Jìn dialect. And according to some informants in my interviews, as well as my own experience, Hū Pǔ speakers are likely to show style-shifting in this variable in different scenarios or with different interlocutors (see the detailed discussion in section 7.4). Therefore, another

linguistic variable was selected in this study to further investigate the interplay between attitudes and awareness.

4.4.1.2 The fricative variable

In order to explore whether explicit awareness is a prerequisite for the attitude-language correlation, a linguistic variable that is below the speakers' conscious awareness should be examined. A feature related to the initial consonants of l-words is chosen here as such a variable: the fricative variable, which means the insertion of a fricative sound [x], [χ], or [ç], that is, a period of frication, after the plosives [p^h, t^h, k^h] or the glottal fricative [h] (see a more detailed description of this feature in chapter 6). This feature was selected because at the time of the fieldwork, I myself, as a native speaker of Hū Pǔ and a linguist, did not even have explicit awareness of this variable. The feature was found when I was listening to the recordings after completing all the interviews. Moreover, none of the participants mentioned this feature when they were talking about Jìn dialect or Hū Pǔ in the fieldwork. It was also impossible for me to prompt them to produce it because I was not aware of the feature. Therefore, it is plausible to say that this feature is less salient than the stress pattern variable, and speakers at least did not have explicit awareness of it. Examining the effects of speaker attitude on the fricative variable could shed some light on the role of attitudes in relation to awareness or salience.

4.4.2 Linguistic data collection: word elicitation task

Collecting speakers' l-words production data was not as easy as examining other segmental variables, because l-words do not appear frequently in natural speech, but are used only in certain specific contexts – a speaker could possibly talk for an hour without using any l-words. Moreover, reading from a wordlist was not a practical task either, because many l-words were known as only existing in colloquial, unwritten, forms. Even if they are managed to be written down, the characters selected would largely influence speakers' pronunciation especially the stress pattern choice. Given this, an l-word elicitation task was

designed in which the informants were encouraged to produce l-words in a natural manner. Sixteen l-words were selected as target words (see Appendix A). In the elicitation task, participants were asked to describe a set of pictures or animations that indicate these 16 target l-words. Still images and photos were used to elicit words that were adjectives and nouns, like 轱辘/*kuəʔ⁴³lu³⁵*/ “wheels” and 窟窿/*k^huəʔ⁴³luŋ³¹*/ “holes”. However, many target l-words were verbs used in very specific situations, such as 趿拉/*t^həʔ⁴³la³¹*/ “to wear shoes like slippers”, or 划拉/*xuəʔ⁴³la³⁵*/ “to scribble”. These verbs were difficult to display using still images. To solve this problem, volunteers were video recorded acting out these particular actions, and the videos were used in the elicitation task. A still from a video is given in Figure 4.2, in which speakers were asked to describe the difference between the two videos. They were very likely to say that the person in the left picture was drawing an apple, while the other person was scribbling. In this way, they produced the target l-word 划拉/*xuəʔ⁴³la³⁵*/ “to scribble”. Also, most pictures and videos were given in pairs, in which the main settings were almost the same but the only difference between the two pictures highlighted the target l-words. This was to help the participants quickly focus on the key point of each picture, so that they would not spend too much time describing the unnecessary background of the pictures.

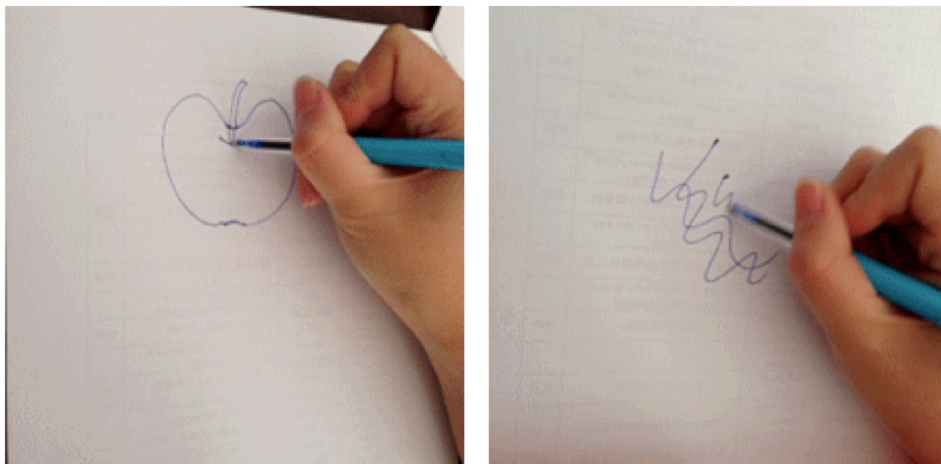


Figure 4.2 An example of the word elicitation task with the target l-word 划拉/*xuəʔ⁴³la³⁵*/ “to scribble”. Note that these were animations in the actual elicitation task.

Altogether 20 sets of pictures and videos were presented to the speakers in PowerPoint slides on a laptop, with one set on each slide (see Appendix B for the full slides used in the elicitation task). Some of the target l-words were elicited more than once in different slides, such as the word 耷拉 /təʔ⁴³la³¹/ ‘droop, hanging’, which was likely to be produced in either Slide 9 or Slide 15 or both. Many other slides were also designed to give participants opportunities to produce the same l-word several times. For example, to describe the difference between the two animations in Slide 5 (see Figure 4.3), a participant is likely to say that in the left picture, the man is carrying a bottle of wine in one hand and (carrying) a basket in the other hand, while the man on the right picture is carrying only a basket. It is possible to elicit the target l-word /t^hiaʔ⁴³liu³¹/ 提溜 ‘to carry something randomly’ three times in this single slide. No slides were included as fillers, but some pictures containing no potential l-words were used to highlight the difference between the pictures in certain slides (e.g., the apple drawing picture in Figure 4.2).



Figure 4.3 Slide 5 of the word elicitation task with the target l-word 提溜 /t^hiaʔ⁴³liu³¹/ ‘carry something randomly’.

As the first task in each data collection interview, the word elicitation task worked exceedingly well not only in collecting target l-words from speakers, but also built a natural and relaxed atmosphere for the whole interview. Participants’ attention was quickly attracted by the humorous videos in which someone tripped up or fell off the bed. And even

for those participants who were illiterate or those who had bad eyesight, this task still worked well.

During the word elicitation task, I tried to keep my role to a minimum and asked the participants to describe the pictures at their own pace. If they did not produce the target l-word, I would encourage them to describe the picture in a different way. But if they still failed, they were asked to pass and go to the next slide. After they finished describing all the pictures, I would go through some of the pictures with the speakers again, and ask them to talk about their use of those l-words. For interviews in group settings, speakers were asked to conduct the task separately one by one⁸, and after they had all completed the task, they were interviewed together about their use of l-words and stress patterns. Most participants naturally produced a good number of target l-words, and occasionally some speakers could even figure out which word I was trying to elicit. Interestingly, many of them also used some non-target l-words, probably because some l-words could be used interchangeably in certain contexts. Therefore, although the elicitation game was designed for 16 target words, participants actually used 38 different l-words in the task (see Appendix A), but for the words brought up by the speakers, there were only a few tokens of each word.

As a result, a total of 4318 tokens of l-words were collected, with an average of 64 tokens per speaker. The majority of the tokens were from the word elicitation task, and some were collected from the surrounding conversation where participants were discussing these words in free speech, but the tokens produced when the speakers were obviously making direct comments on the stress pattern variable, or imitating others, or repeating what I was saying were removed from the data set. Another set of tokens was also removed because they were not clearly pronounced. The remaining 3566 tokens were transcribed and further analyzed on stress pattern and fricative variation, which will be detailed in Chapters 5 and 6.

⁸ Usually when one speaker is describing the pictures, I ask the other speakers to complete the attitudinal questionnaires on their own, in order to avoid mutual influence on linguistic production.

4.5 Attitudinal index scores

Another essential part of this study is to collect and analyze speakers' attitudinal information. As previously mentioned, this dissertation collects speakers' overt attitudes using questionnaires. This section will elaborate on the methodology concerning the attitude data.

4.5.1 The questionnaire: AAS

Oppenheim (1992; as cited in Redinger, 2010) described the measurement of attitudes as an attempt to “place a person's attitude on the straight line or linear continuum in such a way that it can be described as mildly positive, strongly negative and so on” (p. 99). Motivated by this, the most common application of Oppenheim's assumption is the use of Likert scales. However, the limitations associated with Likert scales have been discussed at length (e.g., Bard, Robertson, & Sorace, 1996; Redinger, 2010; Llamas & Watt, 2014; Watson & Clark, 2015). For example, Llamas and Watt (2014) mentioned four disadvantages of using Likert scales: the ambiguity of the mid-point, the limitations imposed on responses, the ‘central tendency’ bias, and the restriction of yielding only ordinal data.

Due to the known flaws of Likert scales, Bard et al. (1996) introduced the methodology of magnitude estimation from psychophysics to linguistic data, in order to gain a more detailed measurement of impressions than ordinal scales. Magnitude estimation requires the participants to assign numerical values to a series of stimuli to reflect the relationship between different stimuli. For example, a participant may give stimulus A an arbitrary value of 2 for intensity, Stimulus B which is perceived as being ten times stronger than stimulus A should be assigned a value of 20. This method does not restrict the participants to a pre-determined range of numbers which can be assigned to each stimulus, and provides measurements of impressions on a numerical scale. However, as Redinger (2010) proposed, matching numbers to stimuli could be a “time-consuming and challenging task” for certain target groups such as school children (p. 105). Therefore, by adapting magnitude

estimation, Redinger developed the magnitude continuum for language attitude research, which allowed participants to mark anywhere on a horizontal straight line to show to what extent they agree or disagree with certain attitudinal statements. This technique, according to Redinger (2010), not only provides “greater freedom of expression” to the participant, but also allows the analyst to gain a fine-grained measurement of the informants’ evaluation (p. 106). More recently, Llamas and Watt (2014) further developed this technique as the Attitude Analog Scale (hereafter, AAS), and proposed that the innovative use of this methodology could be helpful in gaining a more detailed picture of variation in attitudes, thus providing a new route to understand the links between attitudes and linguistic behavior in large-scale sociophonetic studies (p. 616).

To adapt this technique to the Hohhot context, written questionnaires containing a set of statements were designed to elicit attitude data from the participants. The statements were about people’s attitudes towards the local community or dialect, like “Old Town people are vulgar” and “It would be sad if Jìn dialect disappeared in the future”. Questionnaires designed for the New Town participants and Old Town participants were also different according to the different issues of interest. The New Town questionnaire contains 23 statements. Apart from speakers’ attitudes towards the Old Town, Jìn dialect and Hohhot, it also investigates speakers’ identities as state-sponsored migrants. The Old Town questionnaire has a different set of 22 statements. The questions concerning attitudes towards the Old Town were changed or reworded because it is not appropriate to ask an Old

Town participant whether “Old Town people are vulgar”⁹. Instead, their identities of being an Old Towner were investigated through statements like “I’m proud of being an Old Towner”, and questions concerning their opinions about the New Town migrants were also added. Many of the statements in these questionnaires were designed on the basis of the authentic opinions generated from the online forum posts as mentioned in section 3.3. For example, the statement “If my child is seeing or dating someone from the Old Town, I would oppose it” is inspired by the posts indicating some people’s ruling out Old Towners or Jìn speakers in their partner selection criterion. Other statements were created according to my personal experience and knowledge of the local contexts (see Appendix C for the full version of the questionnaires). Designing these statements according to the very specific local contexts proved to be crucial in detecting the genuine overt attitudes of the participants, because these questions are also the actual issues they are facing in their real life, like whether they will be happy if their daughter marries a Jìn speaker, and whether they will move to the Old Town if they have a new apartment there. Some informants even brought up their own experience of how they made these choices in real life. This allows the participants to take serious consideration of the questions instead of giving an answer simply to meet social desirability.

⁹ In Hohhot, the Old Town people are very sensitive when talking about the long-standing prejudice against the Old Town. So in the interviews with Old Town participants, I was often inclined to show a neutral or positive attitudes towards the Old Town, and many participants would start talking about the prejudice issue by themselves in natural conversation. However, if this issue was brought up by me, it might have hurt their feelings and they might be reluctant to talk about it. Therefore, the statements in the Old Town AAS questionnaire were also designed to be more neutral or positive than those in the New Town questionnaires.

16. 如果我孩子的男/女朋友是旧城人，我会反对他们交往。

If my child is seeing or dating someone from the Old Town, I would oppose it.

同意 I agree _____ 不同意 I disagree

17. 此地话很幽默。

Jin dialect is humorous.

同意 I agree _____ 不同意 I disagree

18. 我今后会一直在呼市生活和工作。

I will live and work in Hohhot in the future.

同意 I agree _____ 不同意 I disagree

Figure 4.4 An example of the AAS questionnaire (see the full version in Appendix C).

An example of part of the questionnaire is shown in Figure 4.4, in which informants were asked to mark on the linear scale to show their degree of agreement with each statement. All the questionnaires were completed during the interview so that I could provide an explanation about the statements as well as the technique if needed, and the participants also had the chance to elaborate on their answers. This helped to increase the validity of people's responses in that they could not mark randomly without truly understanding the statements, and I myself could make sure that they did not put the mark on the wrong side by error. For older participants who could not physically complete the questionnaires by themselves due to bad eyesight or illiteracy, I read the questions and asked them to describe their opinions on each statement, and then marked for them accordingly. In this case, I did the markings in front of them so that they could express disagreement if they felt my markings were not appropriate. In the group interviews, speakers were usually asked to complete the questionnaires separately first, and then we talked about the questions together. However, in a few group interviews with older participants who could not complete the questionnaire themselves, I had to read the questions and asked them to give their responses one after another, which might have led to some mutual influence of their attitudes between the group members (see section 5.5 for more discussion of this issue). As a result, all participants successfully completed a valid attitudinal questionnaire in the fieldwork.

After the questionnaire, a semi-structured interview was conducted to further understand the speakers' beliefs and thoughts about the communities and dialects in Hohhot (see Appendix D for the interview topics and questions). In the actual fieldwork, I managed to combine this part with the questionnaire completion part to some extent in most of the interviews - when the informants were elaborating on their answers to the AAS questionnaire, it was often very easy and natural to lead the discussion to related topics in the scheduled interview. For example, when talking about the statement "I love living in Hohhot" in the AAS, participants often naturally began to talk about why they liked or disliked Hohhot. And this was always the best time for me to bring up related questions from the interview topics like "What are the best and worst things about growing up and living in Hohhot?" and "If an outsider was complaining about Hohhot, would you defend it, even if you agreed with what they said?" This method was proved to be very successful in my fieldwork. On the one hand, it made the attitudes data collection procedure more condensed and to the point. Speakers would seldom feel tired or bored because of answering questions on a similar topic repetitively. On the other hand, the in-depth discussion on each statement of the AAS further increased the validity of the speakers' responses. Speakers had plenty of time to think about their true opinions and attitudes during the discussion, and there were actually several cases where the speaker had a second thought about their answer after discussing the related topics and changed their marking.

4.5.2 Calculating attitudinal index scores: Principal Component Analysis

After collecting the questionnaire responses, the length of the participants' marking on each linear scale was measured in millimeters and then the percentage of this length out of the full scale length was logged. All the responses on the AAS were examined and adjusted so that higher scores always represented more positive attitudes.

To analyze the questionnaire responses data, the technique of principal component analysis (hereafter, PCA) was used, which has been applied broadly in analyzing psychological

measurements using questionnaires, especially in attitude studies (see e.g., Akay & Toraman, 2015; Borenic, 2011; Cybulska & Borenic, 2011; McKenzie, 2006; Schilling, 2013).

PCA is a statistical method of dimension reduction, through which the original set of variables is transformed into a substantially smaller set of uncorrelated variables without losing too much information in the original data set (Duntelman, 1989). The purpose of using PCA in this study is two-fold: 1) The attitude questionnaire contains 23 statements (22 in the case of the Old Town questionnaire), so the participants' responses form a data set of 23 variables. PCA reduces these variables to a more manageable size, so that they are easier to interpret in further analysis. 2) The PCA process also helps to explore the underlying structure of the questionnaire responses, so that we can understand which group of questions may cluster together, potentially indexing the same underlying attitudinal aspects.

Since the AAS questionnaires for the New Town and Old Town participants are composed of different statements, they are also analyzed separately.

4.5.2.1 PCA of New Town AAS

Before running PCA on the New Town responses, the original data set was analyzed in R (R Core Team, 2015) to identify whether it was appropriate for factor analysis. To do this, first, the correlation matrix formed by participants' responses to the 23 questions was checked to make sure that all the variables correlated fairly well. Two of the questions (Q18, Q15) were eliminated because they had little correlation with other variables (Field et al., 2012). Then, Bartlett's test of sphericity was run using the *cortest.bartlett()* function from the *psych* package (Revelle, 2015) to check whether the correlations between all the variables were sufficiently large for PCA. The Bartlett's test result of the current matrix was highly significant, $\chi^2(210) = 395.144$, $p < 0.001$, so PCA was deemed appropriate.

PCA was carried out using the *principal()* function in the *psych* package in R. The method of oblique rotation (oblimin) was also used to give a better picture of the variables' loading on

each factor. The result of PCA revealed four main factors, which in combination explained 60% of the total variance. We can see this by examining the scree plot shown in Figure 4.5, which plots the loadings (y) against the factor number (x) (Baayen, 2008; Field et al., 2012; Revelle, 2016). The cutoff point should be a sudden drop or a clear discontinuity before the points reach a stable plateau, which is shown by the circle. However, the scree plot method is sometimes accused of being subjective because it may lead to different interpretation (Revelle, 2016). Therefore, several other statistical criteria were adopted to make decisions on how many factors to retain, including parallel analysis, Minimum Average Partial criterion (MAP), and Very Simple Structure Criterion (VSS)¹⁰ (see Chapter 5.4 in Revelle, 2016 for detailed descriptions of these methods). The results of these criteria all confirmed the decision of taking four factors.

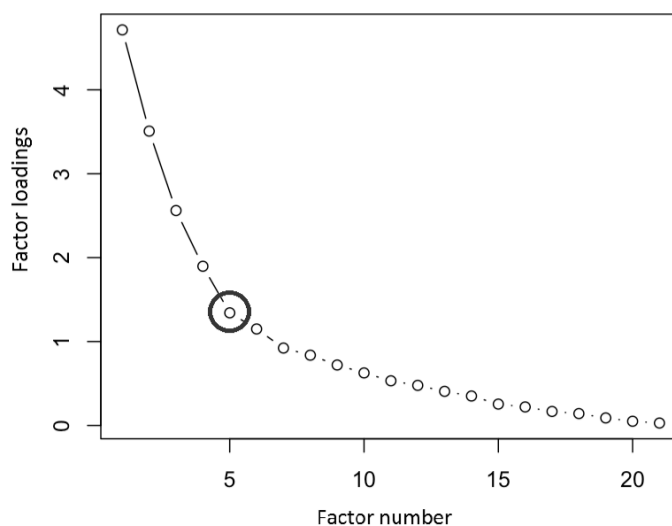


Figure 4.5 Scree plot from PCA of the New Town AAS.

¹⁰ To put it simply, the VSS procedure sees the extracted factor matrix as having a simpler structure than the original matrix and examines how well the simplified matrix reproduces the original matrix. The MAP will extract factors until the average squared partial correlation is minimized. Parallel analysis creates a random parallel data set of the same size as the real data and extracts factors until the eigenvalues of the real data are less than those of the parallel data. For detailed descriptions, see chapter 5.4 in Revelle (2016).

Table 4.3 Four factors revealed by PCA of the New Town AAS and the questions that have higher loadings on each factor (only loadings over .40 were shown).

Questions	Factors			
	1	2	3	4
Q19	0.72			
Q14	0.7			
Q2	0.7			
Q13	0.63			
Q3	0.62			
Q17	0.58			
Q20	0.58			
Q1	0.55			
Q6		0.94		
Q5		0.81		
Q9		0.78		
Q16			0.76	
Q7			0.72	
Q4			0.72	
Q11			0.72	
Q8			-0.6	
Q23				0.86
Q10				0.85
Q22				0.66
Q12				0.42

Table 4.3 shows the results of four factors revealed by PCA, in which the numbers indicate how much each question can contribute to each factor. The four factors or principal components revealed by PCA are shown in Table 4.4. The questions included in each factor can be used to explain what this factor actually means. For example, questions in the first principal component (PC1) all seem to relate to people's emotional or rational attitudes towards Jìn dialect, so PC1 was labeled as "attitudes towards Jìn dialect" or ATTJIN.

Table 4.4 Four principal components revealed by PCA of New Town AAS.

PC1(ATTJIN) 'attitudes to Jìn dialect'	Q19	It would be sad if Jìn dialect disappeared in the future.
	Q14	Learning to speak Jìn dialect is very helpful.
	Q2	Learning to speak Jìn dialect is very interesting.
	Q13	Jìn dialect is vulgar.
	Q3	People can easily get by in Hohhot without knowing any Jìn dialect.
	Q17	Jìn dialect is humorous.
	Q20	There should be some news reports in Jìn dialect on TV and broadcasting.
	Q1	Jìn dialect sounds pleasant.
PC2(IDMIG) 'migrant identity'	Q6	I'll feel close and intimate if I know someone is also a state-sponsored migrant.
	Q5	I am proud of being a state-sponsored migrant.
	Q9	When people ask me where I am from, I often mention I'm a migrant.
PC3(ATTOT) 'attitudes to Old Town'	Q16	If my child is seeing or dating someone from the Old Town, I would oppose it.
	Q7	Old Town people are vulgar.
	Q4	If I find a job in the Old Town, I'll move there.
	Q11	Except for occasional shopping or touring, I seldom go to the Old Town.
	Q8	The local Jìn dialect can represent Hohhot culture.
PC4(HOHORE) 'future orientation'	Q23	If I have an opportunity to live in another city that is similar to Hohhot in economic developments, I will stay in Hohhot.
	Q10	I love living in Hohhot.
	Q22	I hope my descendants can live and work in Hohhot in the future.
	Q12	If I have an opportunity to live in bigger cities like Běijīng or Shànghǎi, I will choose to go.

The second component PC2 was labeled as “state-sponsored migrant identity” or IDMIG since the three questions under this factor (Q6, Q5, Q9) are all related to how much the speakers emphasize their identity as state-sponsored migrants.

In the same way, most questions contributing highly on PC3 were found to connect to people's attitudes towards the Old Town and the local community. Thus, PC3 was labeled as "attitudes towards the Old Town" or ATTOT.

Questions Q23, Q10, Q22, Q12 were clustered as the fourth factor, which related to whether people would like to stay in Hohhot in the future, and how strongly they were attached to Hohhot. So this factor was labeled as "future orientation" or HOHORE.

Table 4.5 Four factor scores revealed by PCA of New Town AAS¹¹.

No.	Speaker	ATTJIN(PC1)	IDMIG(PC2)	ATTOT(PC3)	HOHORE(PC4)
1	NOM1 ¹²	-1.37601	-2.49214	-0.49492	-0.86082
2	NOM2	1.408693	0.821118	-0.569	-0.33241
3	NOM3	1.415084	0.338748	1.158277	-1.00561
4	NOM4	0.11272	-0.91978	0.017754	-0.52859
5	NOM5	1.076835	0.74987	0.761199	-0.90444
...
34	NYF6	0.01835	-0.70607	0.870941	-0.02725
35	NYF7	0.244372	0.013804	-0.51703	1.248484

So as the result of PCA, the 23 variables/questions in the original data set were reduced to four subscales or factors. By adding the *score=TRUE* command to the final PCA model, the following four attitudinal scores were then assigned to each speaker, which corresponded to their attitudes concerning the four factors (see Table 4.5):

¹¹ The attitudinal scores for all New Town informants can be found in Appendix F.

¹² NOM1 is a speaker code where the first letter N indicates that the speaker is from the New Town. The second letter O means that the age group of the speaker is the older group. The third letter M indicates the sex of the speaker to be male. The number means that he is the first speaker in the New Town older male group. All other codes for speakers in this thesis can be interpreted in the same way.

ATTJIN (PC1): attitudes to Jìn dialect

IDMIG (PC2): migrant identity

ATTOT (PC3): attitudes to the Old Town and Old Town people

HOHORE (PC4): future orientation, whether to stay in Hohhot in future

These scores can be used to assess the relative attitudes of one person compared to another: a higher number means that the speaker has more positive attitudes.

4.5.2.2 PCA of Old Town AAS

A similar process was conducted on the Old Town questionnaire. 32 speakers from the Old Town were interviewed in the fieldwork, but one speaker OMM1's questionnaire responses were removed from the data set because I used the wrong type of questionnaire in his interview¹³. Therefore, the Old Town questionnaire analysis introduced in this section is based on 31 speakers.

For PCA to work, the correlation matrix formed by the responses of the 22 questions was examined first to make sure it was appropriate for the analysis. Six questions (Q4, Q10, Q14, Q8, Q18, Q7) were excluded because they had little correlation with other questions. Q9 (If I am talking to another Jìn speaker, I will definitely speak Jìn dialect) was also removed, because during the interview, I found that only bi-dialectal speakers could answer it in a valid way. The new matrix formed by the responses of the remaining 15 questions was

¹³ OMM1's parents were state-sponsored migrants, and he was interviewed together with his wife who was also a second-generation migrant in the New Town. So I used the New Town questionnaire in his interview. But later when I found out that he was actually born and raised in the Old Town, it was already a situation where it would be too embarrassing to change the questionnaire to the Old Town one as it could emphasize the New Town/Old Town identity distinctions between him and his wife.

examined by the Bartlett's test, and the result was highly significant, $\chi^2(105)= 206.9934$, $p<0.000$, so PCA was deemed appropriate.

Again, PCA was carried out using the the *principal()* function with oblique rotation. The scree plot (see Figure 4.6) shows that either taking one or four factors is interpretable. After taking the results of other statistical criteria (VSS, MAP, and parallel analysis) into consideration, one factor was extracted from the Old Town AAS responses, which explained 72% of the total variance.

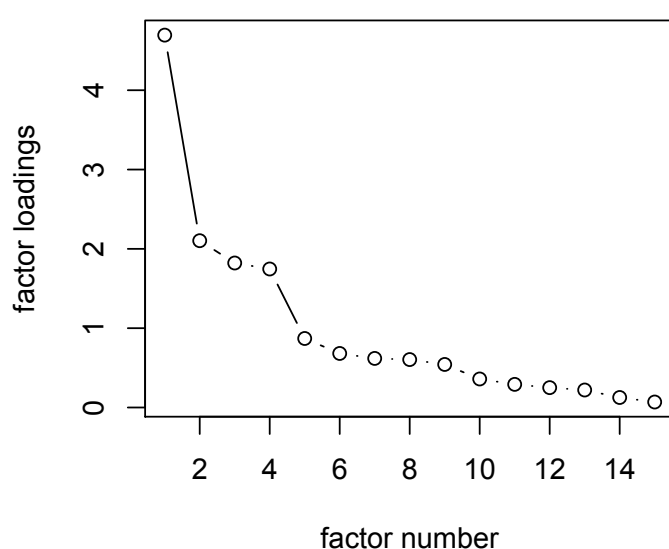


Figure 4.6 Scree plot from PCA of Old Town AAS.

Questions that have high loadings on this factor are listed in Table 4.6, which means that the variation in Old Town speakers' attitudes are better revealed in these questions. Most of them are related to people's attitudes towards Jìn dialect. Then, by adding the *score=TRUE* command to the final PCA model, we achieved a single attitudinal index score ATT for each speaker, as is shown in Table 4.7 (see Appendix F for all Old Town informants' ATT scores). Again, a higher score means that the speaker holds more positive attitudes towards the local dialect and emphasizes their Old Town identity more.

Table 4.6 Questions that have high loadings on the one factor extracted from the Old Town AAS responses (only questions with loadings higher than 0.4 are presented.)

Loadings	Questions	
0.78	Q19	There should be some news reports in Jìn dialect on TV and broadcasting.
0.77	Q1	Jìn dialect sounds pleasant.
0.74	Q11	If my children speak Jìn dialect, I would oppose it.
0.74	Q17	The local Jìn dialect can represent Hohhot culture.
0.72	Q5	I'm proud of being an Old Towner.
0.63	Q21	I hope my descendants can live and work in Hohhot in the future.
0.60	Q16	I hope to see my children and grandchildren speak local Jìn dialect.
0.56	Q12	It would be sad if Jìn dialect disappeared in the future.
0.49	Q13	The reconstruction in the Old Town has destroyed the culture of Hohhot.
0.45	Q20	There should be some comedies in Jìn dialect on TV and broadcasting.
0.40	Q6	Immigrants in the New Town are also local.

Table 4.7 Factor scores revealed by PCA of Old Town AAS.

NO	Speaker	ATT (PC1)
1	OOM1	0.84148899
2	OOM2	1.29603851
3	OOM3	-1.70765941
4	OOM4	-0.81969155
5	OOM5	0.25654084
...
31	OYF6	0.11884216

Therefore, the process of PCA calculated four attitudinal index scores for each New Town speaker (ATTJIN, IDMIG, ATTOT, HOHORE), and one attitudinal score for each Old Town speaker (ATT). These scores formed several numeric variables that represent different aspects of speakers' socio-psychological attitudes. They do not simply define participants' attitudes in terms of positive, negative or neutral, but regard attitude as a gradient concept, which provides a new perspective to explore the attitude-language correlation.

4.6 Social interaction scores

4.6.1 Collecting social networks data

The last part of the data collection phase was to collect speakers' social networks information. The method of relating language variation to speakers' social networks structure quantitatively was first systematically applied by Milroy (1980) and looked at a person's social network as "a series of links which spread out through a society, linking people to one another" (Fox, 2007). Existing works have presented various methods of collecting quantitative data on individuals' social networks, however, many of them involved ethnographic-oriented data collection procedures (Fox, 2007) such as participation observation (e.g., Milroy, 1980; Bortoni-Ricardo, 1985; Li, 1994; Clark & Trousdale, 2013), which were not applicable in this study due to the limit of the total fieldwork time and the interview time for each speaker.

Therefore, a "wedding invitation task" was developed, in which the informants were instructed to make an imaginary wedding guest list. The original design of this task is shown in Appendix E. Participants were asked to list the people they were going to invite, and give information about who were their common and emergency contacts, their relationship to each contact, and whether the contacts were Jin speakers, etc. This task allows the speakers to self-report their own social networks within a short period of time,

and it does not sound too intrusive because planning a wedding list is a very common practice in Hohhot. This task was given to the participants at the end of the interview, and they were asked to finish and return it to me after a week. To protect the participants' personal information, I asked them to cut off the 'guest name' column before returning the list.

In the fieldwork, however, this wedding invitation task did not work as well as I thought. Few participants returned their responses to me during the early weeks. Some claimed that the paper was lost, or simply gave no responses. Some were a bit exhausted after the long interview and seemed reluctant to do this especially after hearing the complex instructions. The few people who completed the task gave a list of only a small number of guests, which could not be a good reflection of their real social networks. Apart from these, another unexpected issue was found with this task: the older participants were over 70 years old at the time of the interview, so many of their friends they would have invited had already passed away, which made the task less meaningful for them.

"Are your parents Jìn speakers?"

"Does your husband/wife speak Jìn dialect?"

"How many relatives do you have here in Hohhot? How many of them are Jìn dialect speakers?"

"Do you have more contact with Jìn dialect speakers or Pǔtōnghuà speakers at your working place/at school/among your friends?"

"Do you have experience of living in a Jìn-dialect-dominant environment for more than a year?"

Figure 4.7 List of questions used in the interview to collect speakers' social contact information.

Therefore, I decided to change the strategy of data collection in this part by asking the participants directly and briefly about their social interaction with Jìn speakers and made brief notes of them in the final part of the data collection procedure. The main questions

used are shown in Figure 4.7, which effectively helped me to obtain speakers' basic social networks information, and the procedure usually took less than 10 minutes.

4.6.2 Calculating the social interaction index score

To extract quantitative data, a calculation chart was created to measure speakers' social contacts with Jìn speakers as shown in Table 4.8. Each speaker was given scores for their degree of integration into the local community in different spheres of their lives, like family, workplace, school, etc. A score of "2" means that the speaker has contact with more Jìn-speakers than non-Jìn-speakers in this particular setting. "0" indicates more non-Jìn-speakers than Jìn-speakers, while "1" stands for roughly equal contact with both sides. In the interview, many participants claimed that they were influenced by their spouse's speech because they had spent more time together. Therefore, a speaker gained "2" points if his/her spouse was a Jìn-speaker. Some informants also mentioned that they had experience living in a Jìn-speaking environment for a certain period of time, which had influenced their knowledge of the local dialect. So another "1" point was added if a speaker had this sort of experience. Speakers who did not have a spouse yet or who were still students were given an "NA" in their corresponding settings "spouse" and "work". Then, the percentages of Jìn speakers in all these settings were calculated for each speaker to be used as their social networks index: SOCNET (see Table 4.8), which represents their social interaction with Jìn speakers. A higher number means that they have more social contact with the local community.

Table 4.8 The calculation chart for speakers' social contact with Jìn speakers¹⁴.

Speaker	parents	spouse	family	work	school	friends	JìnExp	SOCNET
NOM1	0	0	0	0	0	0	1	0.071
NOM2	2	2	1	2	1	1	1	0.714
NMF6	0	2	1	0	0	0	0	0.214
NYM1	1	NA	2	1	0	0	1	0.417
NYM2	2	NA	2	0	0	0	1	0.417
NYM5	1	0	2	1	0	0	1	0.357
...
NYF1	0	NA	0	NA	0	0	0	0

4.7 Demographic variables

Apart from attitudes and social networks, speakers' demographic background information was also collected in a part of the questionnaires. These traditional socio-demographic variables were also taken into consideration when examining the correlation between individuals' linguistic production and their social characteristics.

As introduced in section 4.2, participants were evenly distributed in age groups, sex, and region (New/Old Town), so these variables were included in the analysis as AGE, SEX and TOWN. Socio-economic status of the participants was investigated in terms of educational background and occupation. Educational background, or EDU, was defined in a binary way according to whether the speaker had received a college education or not. This criterion could differentiate social class in most cases especially for older and middle-aged participants, because before the 1990s, people who graduated from universities were in the

¹⁴ SOCNET scores for all informants can be found in Appendix F.

minority and could easily find a good job as cadres or intellectuals. People who did not go to college were often working class. However, in the third generation, this social division was no longer that obvious because going to college became very common. Occupation information was also included as a binary variable OCCU, in terms of white- or blue-collar. Participants' relationship to me was recorded as RELATION, in which informants were classified as friends, acquaintances or strangers. For the New Town participants, most of them were state-sponsored migrants and their offspring, but some were not, so this factor was also included as a binary variable: ST-SP. For the Old Town participants, their language background was included as a three-level categorical factor LANG according to whether they were Jin speakers, non-Jin speakers, or bi-dialectal. Also, some Old Town informants had already moved to the New Town at the time of the interview, which could influence their attitudes and linguistic production, so the factor MOVNT was included in the analysis concerning the Old Town speakers.

4.8 Quantitative analysis techniques

The main statistical technique used in this thesis is mixed-effects modeling (Baayen, Davidson, & Bates, 2008). As the two linguistic variables selected are both coded as categorical factors, generalized linear mixed-effects regression models were built, using the *glmer()* functions in the *lme4* library (Bates, Maechler & Bolker, 2011) in the software package R as the primary tool.

Mixed-effects modeling has gained increasing popularity in sociolinguistics in recent years because it allows the researcher to simultaneously consider all factors that potentially contribute to the understanding of the structure of the data (Baayen et al., 2008). Compared to traditional statistical tools like ANOVA, VARBRUL or GoldVarb, mixed-effects models are more flexible, powerful, and provide more accurate results (Johnson, 2009; Baayan, et al., 2008). They also address the concerns of traditional ANOVA by offering better solutions to

dealing with missing data, treating continuous and categorical responses disparately, and other problems posed by repeated observations (Baayen et al., 2008).

Mixed-effects models not only comprise standard fixed-effects factors, which can be replicated and generalized to an entire population such as sex or age, but also take control of random effect factors that arise from random causes, such as the speakers and words selected in the study. Random effects are often not replicable - two studies with the same linguistic phenomenon of interest could both examine men and women as the sex factor, but the speakers they recruited and the words they selected cannot be the same. By including a random intercept effect of *speaker*, the model allows different speakers to have different baselines in the dependent variable, or in the current study, their degree of adopting Jìn features – some speakers are generally more likely to adopt the features than others, regardless of their sex, age, or other social-related factors included in the fixed effects. These idiosyncratic variations of the speaker may not be of interest in the current study, but they are controlled by the model if speaker is included as a random intercept. The random effects structure may also include random slopes, which allow the effects to vary across different speakers/words. For example, the effect of speaker attitude on their adoption of the stress pattern variable might be larger on some l-words than it is for others. Therefore, including random intercepts and slopes increases the statistical power for detecting between-speaker and between-word effects (Bates, Kliegl, Vasishth, & Baayen, 2015).

For models that include multiple fixed and random effects factors, the question of choosing the appropriate random effects structure is often complex. Barr, Levy, Scheepers, and Tily (2013) recommend a maximal random-effects structure in which all possible components should be included. However, recent work by Bates et al. (2015) points out that estimation of maximal models may not converge simply because the model is too complex to be supported by the data, or even under convergence, the results could be uninterpretable due to overparameterization. They advocate for a more parsimonious model in which the complexity of the random effect structure is diagnosed by PCA and non-significant variance

components are dropped. Models in the current study often test many linguistic and social factors at the same time, as well as their interactions, which may lead to complex structures relative to the available data size. Therefore, I chose to follow the parsimonious models as proposed by Bates et al. (2015) in terms of random effects structure.

5 Stress pattern variation

The first linguistic variable examined in this study is the stress pattern of l-words. In this chapter, I explore how speakers' adoption of the stress pattern variable is related to their social characteristics, especially their socio-psychological orientations and social networks.

5.1 Description of the variable

The stress pattern of l-words has been discussed in previous Chinese phonology literature (e.g., Zhao, 1979; Xu, 1981; Wang, 1994), but it has never been explored from a variationist sociolinguistic perspective.

As mentioned in section 4.4.1, l-words display variation between a weak-strong pattern and a strong-weak pattern. In Jìn dialect, the stress pattern of l-words is perceptually more similar to an iambic or weak-strong type. The vowels in the first syllable of l-words are usually reduced and followed by a glottal stop, which is also termed the entering tone, as explained in section 3.1. The entering tone typically makes the syllable sound shorter than normal stress syllables. Therefore, in the pronunciation of l-words in Jìn dialect, the stress of the second syllable outweighs that of the first syllable, thus a weak-strong pattern is perceived.

Table 5.1 Examples of l-words that are pronounced in different stress patterns in Jìn dialect and Pǔtōnghuà (see more examples in the full l-word list in Appendix A).

	Jìn dialect	Pǔtōnghuà
轱辘 ‘wheels’	/kuəŋ ⁴³ lu ³⁵ /	/ku ³⁵ lu ⁰ /
耷拉 ‘droop, hanging’	/təŋ ⁴³ la ³¹ /	/ta ⁵⁵ la ⁰ /
划拉 ‘scribble’	/xuəŋ ⁴³ la ³⁵ /	/xua ³⁵ la ⁰ /
窟窿 ‘holes’	/k ^h uəŋ ⁴³ luŋ ³¹ /	/k ^h u ⁵⁵ luŋ ⁰ /
	weak-strong	strong-weak

However, in modern Pǔtōnghuà, where the entering tone has disappeared (more accurately, merged with other tones), these l-words are pronounced in a sort of trochee or strong-weak stress pattern, where the first syllable is pronounced with a normal tone and stress, but the second syllable is with weak stress. In the weak syllable, neutral tone (Chao, 1968) is often used where “the tone range is flattened to practically zero and the duration is relatively short” (pp. 35-36). For example, in Table 5.1, the word 轱辘 ‘wheels’ is pronounced as /kuəŋ⁴³lu³⁵/ in Jìn dialect where the first syllable ends with a glottal stop with shorter duration. However, in Pǔtōnghuà, this word becomes /ku³⁵lu⁰/, with full tones and stress in the first syllable, but the stress is lost in the second syllable where a neutral tone is adopted. Hū Pǔ speakers often use weak-strong and strong-weak patterns variably, which makes the stress pattern variable a good feature to explore their degree of adopting Jìn features.

5.2. Methodology

L-words tokens collected from the word elicitation task and interviews were further analyzed in terms of their realized stress patterns. As mentioned in section 4.4.1, some l-words are only used in Jìn dialect, which do not have a Pǔtōnghuà form. These l-words were removed from the data set because they can only be produced in weak-strong forms, showing no variation in stress pattern at all. The remaining 2812 valid tokens were hand

coded in Praat (Boersma & Weenink, 2017). Coding decisions were based on auditory perception, combined with visually inspecting the spectrogram of each token.

Most of the tokens were straight-forward in terms of the realization of either a weak-strong pattern (w-s) or a strong-weak pattern (s-w). In the word list (see Appendix A), there are also a few l-words whose Jìn and Pǔtōnghuà forms do not strictly follow the weak-strong/strong-weak distinction, but they are sometimes realized as a strong-strong pattern with two full syllables (e.g., the Pǔtōnghuà form of 陀螺 ‘the spinning top’ is /t^huo³⁵luo³⁵/) or a weak-weak pattern, with the entering tone on the first syllable and a neutral tone on the second (e.g., the Jìn form of 趿拉 ‘to wear shoes like slippers’ can be realized as /t^həʔ⁴³la⁰/). These words are included because they still have a contrast in stress pattern between the Jìn form and the Pǔtōnghuà form. A strong-strong pattern is also more oriented to Pǔtōnghuà, whereas a weak-weak pattern is still considered to be a Jìn feature. So in these cases, the strong-strong pattern was coded as strong-weak (s-w), and the weak-weak pattern was coded as weak-strong (w-s). There were also some tokens produced by the New Town older speakers (first generation migrants), that seemed to be between w-s and s-w due to the influence of their accents from original hometowns. For these tokens, the judgment was made by inspecting other acoustic features on the spectrogram, such as the duration of the vowels in the two syllables, or whether a glottal stop is involved in the first syllable.

To verify the perceptual coding, a subset of the data (about 2000 tokens) was inspected using Praat for acoustic characteristics. For each token, the duration of the vowel in the first syllable and the duration of the entire word was measured. The ratio of the vowel duration to the word duration was then calculated to represent the comparative weight of the first syllable in each token. Mixed-effects models were fitted in R with the stress pattern binary coding as the dependent variable, and this duration ratio was included as the independent variable. The results showed that there was a very clear statistical correlation between the duration ratio and the auditory coding ($p < 0.001$). As shown in the plot in Figure 5.1, tokens with higher duration ratio, whose vowel duration in the first syllable is

comparatively longer, were more likely to be coded as strong-weak patterns rather than weak-strong patterns. The manual coding was therefore justified.

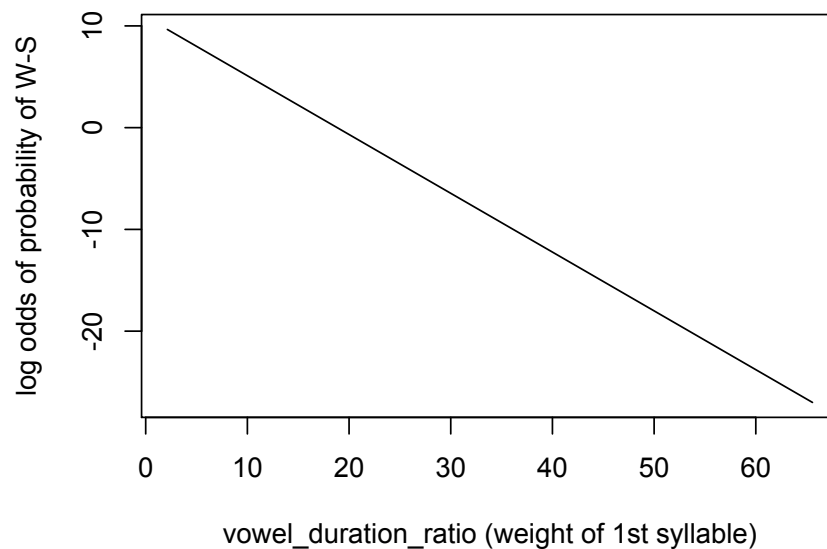


Figure 5.1 Correlation between the vowel duration ratio and the log odds of the probability of a token being coded as weak-strong (w-s).

Next, I explore the social and linguistic constraints operating on the stress pattern variable with the following research questions in mind:

1. In large-scale dialect contact situations, will speakers' linguistic variation be conditioned by their social attitudes? If so, how will speakers' attitudes influence their linguistic production?
2. Are speakers' explicit attitudes collected by overt questionnaires likely to predict linguistic behavior?
3. Will speakers' attitudes play an independent role in language change if their social contact is also taken into consideration?
4. How will attitudes interact with social contact to influence speakers' linguistic production?

5. What is the role of attitudes in koinéization? Will attitudes play different roles in different generations of speakers in a new dialect formation scenario?

6. Are speakers' attitudes and identities important in shaping the outcome of koinéization?

As New Town speakers and Old Town speakers were investigated using different AAS questionnaires, the effects of the attitudinal scores had to be explored in the New Town and Old Town data sets separately. Before doing that, the social and linguistic distribution of the stress pattern variable was first investigated in the combined data set to see the general language change situation of this variable. The following sections present the results from the combined data set, the New Town data set and the Old Town data set.

5.3. Results of the full data set

A binomial mixed effects model was first fitted to the full data using the *glmer()* function in the *lme4* library in R. The dependent variable was the realization of l-words stress pattern being produced as weak-strong, which oriented to the local Jìn-feature. The demographic factors shared by both New Town and Old Town speakers, including AGE, SEX, TOWN, EDU, OCCU, and RELATION, were tested as fixed effects, as well as interactions between these factors.

To avoid multicollinearity in the fixed effects, the statistical diagnostic Variance Inflation Factor (VIF) was used to test whether two or more of these predictors were highly correlated. The primary risk of multicollinearity between predictors is the difficulty in assessing the individual importance of a predictor. To test this, a VIF score is calculated for each predictor, and as Lefcheck (2012) suggests, a VIF = 1 indicates no collinearity, whereas increasingly higher values suggest increasing multicollinearity. The strategy used in this study is to sequentially drop the variable that has the highest VIF, recalculate and repeat

the process until all VIFs are below a certain cutoff (3 as suggested in Zuur et al., 2010). As a result, RELATION was removed from the predictors, probably because it was correlated with AGE and TOWN due to my own social networks – participants who were classified as “friend” were mostly New Town younger group speakers.

The 2812 tokens from 67 speakers in the combined data set were fitted into the regression model using the bobyqa optimizer. The model was hand-fit using a forward stepwise selection technique, and model selection was guided by the Akaike Information Criterion (Akaike, 1974). Each model contains *speaker* and *word* as random intercepts. The social predictors were also tested as random slopes for word if they were included in the fixed effects. The best-fit model had the fixed effects of EDU, and a three-way interaction between TOWN, AGE and SEX. As random effects, the model had intercepts for *speaker* and *word*, as well as by-word random slopes for AGE, TOWN, SEX, and EDU.

Table 5.2 Output of the best model for the stress pattern variable in the full data set.

	Estimate	Std.Error	p value	
(Intercept)	-2.9374	1.1725	0.012239	*
TOWNoldtown	10.2117	1.7009	1.93E-09	***
AGEmiddle	3.4504	1.2175	0.004597	**
AGEyounger	3.3750	1.2172	0.005560	**
SEXmale	3.4703	1.1991	0.003803	**
EDUyes	-1.8384	0.7569	0.015146	*
TOWNoldtown * AGEmiddle	-8.4786	1.9822	1.89E-05	***
TOWNoldtown * AGEyounger	-8.6541	1.9861	1.32E-05	***
TOWNoldtown * SEXmale	-6.6768	2.0204	0.000951	***
AGEmiddle * SEXmale	-3.7703	1.7039	0.026914	*
AGEyounger * SEXmale	-1.8968	1.6104	0.238876	
TOWNoldtown * AGEmiddle * SEXmale	8.7885	2.6316	0.000839	***
TOWNoldtown * AGEyounger * SEXmale	6.4494	2.5678	0.012016	*

The output of the best model is presented in Table 5.2, and the effects of significant factors are plotted in Figures 5.2 and 5.3. In these plots, the Y-axes all represent the log odds¹⁵ of the probability of the stress pattern being realized as weak-strong (w-s, the Jìn feature). A larger number refers to higher probability of being produced as weak-strong. Figure 5.2 shows the effects of EDU on speakers' stress pattern production. Speakers who had received a college education were significantly less likely to use the local weak-strong pattern than those who had not ($p < 0.05$)¹⁶. This result can be expected, because well-educated people have been found in sociolinguistic patterns to use more standard forms (e.g., Labov, 2001).

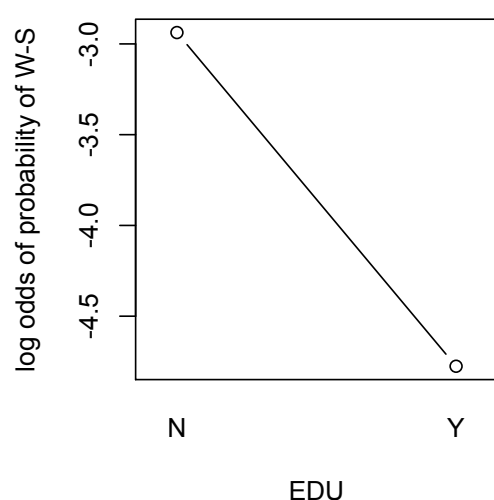


Figure 5.2 Effect of educational background (EDU) on stress pattern variation in the full data set. Y refers to speakers who have received a college education.

¹⁵ Log odds are an alternative way of expressing probability. They are often used in logistic regression models.

¹⁶ All models reported in this dissertation are mixed effects models. Figure 5.2 and other plots in the remaining analysis all present the model predictions. Therefore, raw data such as percentages or token numbers were not labeled in these plots to avoid misinterpretation of the results. These raw data for all speakers are provided in Appendix G. Also, the informants' detailed social information, attitudinal scores and social networks scores can be found in Appendix F.

Plots of the three-way interaction between AGE, SEX and TOWN are shown in Figure 5.3.

The X-axes are the three age groups or generations. The black trend lines refer to Old Town speakers, and the gray lines represent New Town speakers. The left plot shows the effect of interaction between TOWN and AGE for female speakers, whereas the effect for male speakers is shown in the right plot.



Figure 5.3 Plots of the three-way interaction between TOWN, SEX and AGE for the stress pattern variable in the full data set.

The language change pattern for the female speakers (the left plot in Figure 5.3) clearly shows a trend of focusing for the stress pattern variable. The older speakers from the Old Town (the very left point on the black line), who are all Jìn speakers, are most likely to produce the weak-strong pattern because they used this variant almost exclusively when speaking Jìn dialect. By contrast, the older speakers from the New Town (the very left point on the gray line), as the first-generation migrants, were significantly less likely to produce the weak-strong pattern ($p < 0.001$), because this feature is rarely seen in either Pǔtōnghuà or their hometown dialects. In the second and third generations (the middle and right points on both lines), however, there is no longer a significant difference in stress pattern

use between New Town and Old Town speakers – they converge to each other and show a similar probability of using the weak-strong pattern.

For the male speakers (the right plot in Figure 5.3), the pattern is slightly different for their stress pattern use. The older speakers from the Old Town (the very left point on the black line) are still significantly more likely to produce the weak-strong pattern than the first-generation migrants in the New Town (the very left point on the gray line) ($p < 0.01$).

However, unlike the female speakers' pattern, a significant difference between the New Town and Old Town speakers was also found in the second-generation middle-aged speakers ($p < 0.01$), and convergence between the two towns was only found in the third generation. This is probably related to the difference in language use between male and female participants in the Old Town middle-aged group, who were mostly bi-dialectal speakers. In the fieldwork, if the participants claimed that they can speak both Jìn and Pǔtōnghuà, they were asked to speak to me in the way they liked or they were comfortable with. Table 5.3 presents the actual language choice of all the Old Town middle-aged speakers in the interviews (see column "interviewLANG"). Five out of six female participants claimed themselves to be bi-dialectal, but they all chose to speak Hū Pǔ or code-switch between Hū Pǔ and Jìn in the interview. Some of them even felt unable to speak Jìn to me because I was not a Jìn speaker. By contrast, for the male participants, the two bi-dialectal speakers both chose to speak Jìn in the interview even though I was speaking Hū Pǔ. Another two male participants claimed that they were monolingual Jìn speakers, thus they also spoke Jìn in the interview. Speakers will use weak-strong patterns in l-words almost exclusively if they are speaking Jìn dialect, so male speakers demonstrated a significantly higher probability of producing the weak-strong forms than the corresponding New Town age group. However, female middle-aged speakers from the Old Town and the New Town, who were mostly speaking Hū Pǔ in the interview, did not show a significant difference in their stress pattern use. This finding confirms the typical sociolinguistic pattern of the sex factor, in which females are more likely to conform to the standard variety than males. The second-generation speakers in the Old Town are gradually

shifting from Jìn to Hū Pǔ or Pǔtōnghuà, and female speakers are leading the change to converge to the New Town speakers in stress pattern use.

Table 5.3 Language use and language choice of the Old Town middle-aged participants.

SEX	Speaker	LANG	interviewLANG
female	OMF1	bi-dialectal	Hū Pǔ
	OMF2	bi-dialectal	code-switching
	OMF3	bi-dialectal	Hū Pǔ
	OMF4	non-Jìn	Hū Pǔ
	OMF5	bi-dialectal	code-switching
	OMF6	bidialectal	code-switching
male	OMM1	non-Jìn	Hū Pǔ
	OMM2	Jìn	Jìn
	OMM3	bi-dialectal	Jìn
	OMM4	Jìn	Jìn
	OMM5	bi-dialectal	Jìn

Another difference between the patterns of male and female speakers can be seen from Figure 5.3 with respect to the first-generation speakers. For the older participants from the Old Town, male speakers (left point, black line, right plot) were significantly less likely to produce the weak-strong pattern than female speakers (left point, black line, left plot) ($p < 0.05$). This can be explained by the participants' distribution in my data. As mentioned in section 4.1, due to the limit of my own social networks, I did not find enough older participants in the Old Town, but had to interview some retired professors in the New Town who were originally from the Old Town. The five male participants in my data were all found in this way, who used to be professors or engineers working in Gōngdà university, and still lived in the university neighborhood at the time of the interview. Two of them became bi-dialectal speakers in this environment, and they both chose to speak Hū Pǔ in the interview. By comparison, the four female older speakers were all monolingual Jìn speakers, who demonstrated a high probability of using the weak-strong pattern. Therefore, this result might be because the male participants could not be considered as representative speakers of the Old Town older group. However, it was also a fact that, at the

time of their generation, it was more likely for males to receive higher education than females in the Old Town society, because large families at the time often prioritized their educational investment in favor of their sons rather than daughters, especially in less-developed areas (see e.g., Rong & Shi, 2001). So this finding is also likely to reflect the sociolinguistic situation of the Old Town participants to a certain degree.

The New Town older group speakers also presented different patterns for males and females. Men (left point, gray line, right plot) were found to be significantly more likely to adopt the local Jìn feature weak-strong pattern than women (left point, gray line, left plot) ($p < 0.01$). This may be related to the fact that most of the male speakers showed more positive attitudes towards Jìn dialect and willingness to integrate into the local community than females. Many of them also had experience of living and working in Jìn-speaking areas for a certain period of time.

So generally, the New Town and Old Town speakers, through three generations of dialect contact, showed growing convergence to each other on the use of the stress pattern in l-words. The first-generation migrants, with the male speakers leading the trend, gradually integrated into the local Jìn group and began to adopt the weak-strong pattern. At the same time, the Old Town Jìn speakers, especially the well-educated people, began to shift to Pǔtōnghuà or Hū Pǔ. As a result, the middle-aged locally-born residents mostly became bi-dialectal speakers and females were leading the change to converge to the New Town speakers in stress pattern use. In the third generation, the new variety of Hū Pǔ was even more stable and younger speakers from both communities showed a similar probability of using the weak-strong pattern.

The findings also indicate that apart from the demographic factors explored in this model, speakers' adoption of the weak-strong pattern may also be influenced by their social networks or attitudes. So next, I explore the effects of these factors in the New Town and Old Town data sets separately.

5.4. Results of the New Town data set

The New Town data set contains 1529 tokens from 35 speakers. These tokens were fit into a binomial mixed effects model in R, again, with the binary stress pattern coding as the dependent variable. Speaker and word were still included as random intercepts. The four attitudinal index scores: ATTJIN (attitudes to Jìn dialect), HOHORE (future orientation), ATTOT (attitudes towards the Old Town), IDMIG (state-sponsored migrant identity), and social networks index score SOCNET for New Town speakers were tested as fixed effects, as well as other demographic factors including AGE, SEX, EDU, OCCU, RELATION, and ST-SP (whether they were state-sponsored migrants). The VIF diagnostic test was used again to test potential multicollinearity between these social factors. This step is essentially important in the present study, especially in teasing apart attitudes from social interaction, because if the attitudinal scores and social networks scores are strongly correlated, I would not be able to know the independent significance of them. RELATION was first dropped as the variable that had highest VIF. I only had friends in the younger group, and all older speakers were strangers to me. OCCU was then removed since it correlated with EDU – speakers who had received college education were also white-collar in the majority of cases. One of the attitudinal scores IDMIG was also eliminated, because it showed collinearity with AGE. Older speakers tended to have stronger emphasis on their state-sponsored identity than other groups because they were the first generation migrants. The remaining eight predictors all had VIF scores below three.

The effects of the eight factors and the interactions between them were tested in mixed effects models using a forward selection technique. The pairwise model comparisons were still guided by the Akaike Information Criterion. The best-fit model had the fixed effects of two-way interactions between AGE and the three attitudinal scores, two-way interactions between SOCNET and the three attitudinal scores, two-way interaction between SOCNET and AGE, as well as SEX. For random effects, the model had speaker and word as random

intercepts, as well as by-word random slopes for AGE, HOHORE, and ATTOT. The details of the best-fit model are presented in Table 5.4.

Table 5.4 Output of the best model for stress patterns in the New Town data set.

	Estimate	Std.Error	p value	
(Intercept)	-4.22815	0.64637	6.10E-11	***
AGEmiddle	2.53429	0.83194	0.002317	**
AGEyounger	2.77199	0.91606	0.002478	**
SOCNET	11.20956	2.01132	2.50E-08	***
SEXmale	0.60522	0.35152	0.085123	.
ATTJIN	1.4282	0.55754	0.010419	*
HOHORE	0.62665	0.51253	0.22146	
ATTOT	-1.35647	0.53355	0.011012	*
ATTJIN * AGEmiddle	-2.06359	0.598	0.000559	***
ATTJIN * AGEyounger	-2.37569	0.62829	0.000156	***
HOHORE * AGEmiddle	1.46067	0.53854	0.006682	**
HOHORE * AGEyounger	0.06025	0.53809	0.91084	
ATTOT * AGEmiddle	1.57087	0.60822	0.009802	**
ATTOT * AGEyounger	-2.2225	0.64578	0.000578	***
ATTJIN * SOCNET	2.7064	1.29725	0.036956	*
HOHORE * SOCNET	-4.65528	1.42391	0.001078	**
ATTOT * SOCNET	3.84714	1.41836	0.00668	**
AGEmiddle * SOCNET	-3.70578	2.96283	0.211023	
AGEyounger * SOCNET	-6.01072	2.73049	0.027712	*

5.4.1 The effects of attitudes

Figure 5.4 shows the plots of the effects of the three attitudinal scores ATTJIN, HOHORE, and ATTOT when they interact with AGE. The Y-axes are the log odds of the probability of the use of the weak-strong pattern. The X-axes in the three plots represent ATTJIN, HOHORE and ATTOT attitudinal scores respectively, where a higher number always means comparatively more positive attitudes. The three line types in the plots refer to the three age groups or generations. In general, the expected direction of attitude effects is that a positive correlation will be found between the use of the weak-strong pattern and positive

attitudes. That is, speakers with more positive attitudes towards the local dialect and communities will have higher probability of adopting the local weak-strong form.

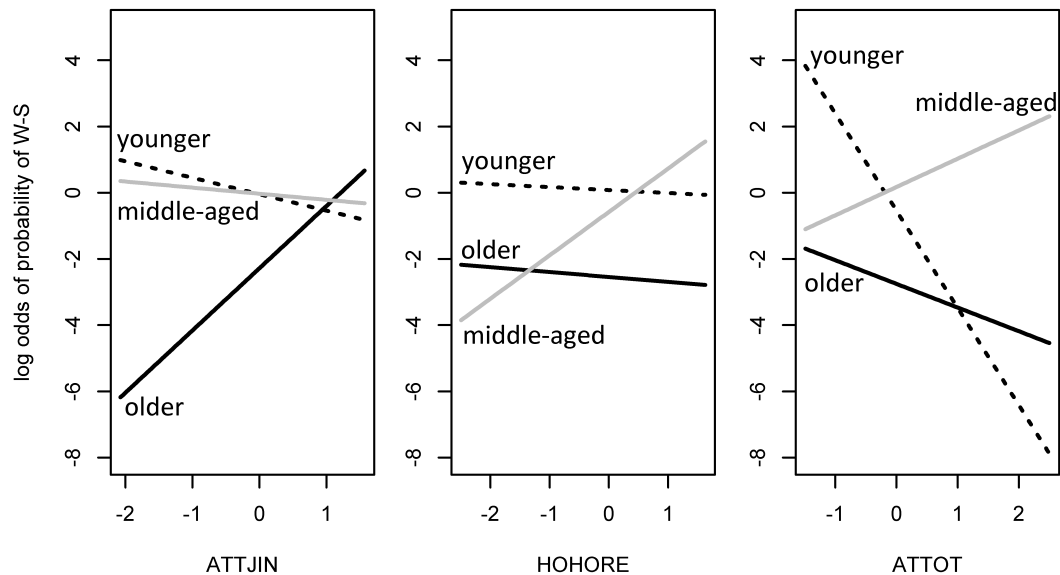


Figure 5.4 Plots of the two-way interactions between AGE and three attitudinal index scores in the New Town data set.

In the older generation, ATTJIN, representing the speakers' attitudes towards Jìn dialect, was found to have significant effects on their stress pattern production ($p < 0.05$). Speakers who demonstrated more positive attitudes towards the local dialect were more likely to adopt the weak-strong pattern, which was as expected. ATTOT, or speakers' attitudes towards the Old Town, was also found to be a significant predictor ($p < 0.05$), however, the direction of the attitude effect was contrary to what I had expected. Speakers with more positive attitudes towards the Old Town were less likely to adopt the Jìn feature: weak-strong pattern. When it came to the middle-aged speakers, their stress pattern choice was predicted by their future orientation score: HOHORE. Those who would like to stay in Hohhot were significantly more likely to use the weak-strong pattern ($p < 0.001$). In the younger generation, however, the pattern became quite different. ATTOT had a very significant effect ($p < 0.001$), but was negatively correlated with speakers' weak-strong pattern use (see the dashed line representing younger speakers in the very right plot).

ATTJIN was also a significant predictor ($p < 0.05$), and the direction of the effect was found to be a negative correlation as well.

To summarize, in the New Town data set, all three attitudinal index scores were found to be significantly correlated with speakers' adoption of the weak-strong stress pattern in l-words. However, the effects of attitudes were different in the three different generations. ATTJIN and ATTOT were significant predictors for the older and younger generations, and HOHORE had an effect in the middle-aged speakers. The direction of the attitudes effects was not always as expected, and some of the attitudinal factors showed negative correlations with speakers' stress pattern use. These attitudes effects found in different generations of the migrant community are closely related to the specific social and linguistic contexts of Hohhot. In the next paragraphs, I interpret these findings with respect to the changes in the social situations of Hohhot and provide some evidence from the sociolinguistic interview.

The older group speakers were the first generation of migrants in New Town. Before they arrived in Hohhot, they had never heard of these l-words being pronounced with a weak-strong stress pattern because this is a unique feature of Jìn dialect. Therefore, it is very likely that they would relate this linguistic feature only to Jìn dialect, and whether or not they choose to adopt this feature may partly depend on their attitudes towards the local dialect, which explains why ATTJIN was positively correlated with their stress pattern use. In the interviews, the older speakers also showed that they had a clear idea of the distinction between Jìn dialect and Pǔtōnghuà, as well as between strong-weak and weak-strong stress patterns. For example, in the excerpt in (5-1), Speaker NOM5 talks about what the weak-strong pattern means to him and why he would sometimes use this feature.

(5-1) NOM5, older group, male

NOM5: (describing the picture of a gourd) This is called /xu³⁵ lu⁰/(s-w).

XW¹⁷: Do you sometimes say /xuə⁴³ lu⁵¹/(w-s)?

NOM5: (laugh) Local people use /xuə⁴³ lu⁵¹/¹⁸(w-s). It sounds ... a bit different... but it's the same word in Jìn dialect.

NOM5: I've been living in Hohhot for such a long time, so I also use some local Jìn dialect. This will help me to show more intimacy with the local people. If you always speak standard Pǔtōnghuà, it's like you are not able to integrate into the local environment. So now I'm speaking a sort of mixed dialect. Like when I'm teaching, I usually use some not very standard Jìn dialect words, then my students will be more interested in my lecture.

ATTOT was also correlated with older speakers' stress pattern production, but the effect was in the opposite direction. This may be related to the fact that some older participants who had higher scores in ATTOT actually had little contact with Jìn speakers in their social circle – some of them had SOCNET scores of zero. Therefore, even though they demonstrated more positive attitudes, they had almost no opportunities to “learn” to use this Jìn feature. This will be further explained in the next section when I discuss the effects of interaction between attitudes and social networks.

For the middle-aged speakers, the attitudinal score HOHORE, representing their future orientation had a significant effect. Speakers who would like to stay in Hohhot in future were more likely to adopt the weak-strong pattern. In the AAS questionnaires, the four questions which contributed to the HOHORE factor seemed to reflect middle-aged speakers'

¹⁷ XW refers to the interviewer: myself.

¹⁸ This weak-strong token was not included in the analysis because it is obviously a direct comment on the stress pattern variable, and the participant was saying that this was used by the local people, not him. Tokens like this were labeled during the coding process and were excluded in the data analysis, as mentioned in section 4.4.2.

emotional attachment to Hohhot to a large extent. Some of them were very locally-oriented, insisting that they would love to stay in Hohhot and expected their children to live here as well. These speakers often spoke highly of Hohhot and presented a deep love for the city as their hometown. For example, in excerpt (5-2), speaker NMF1 expressed her emotional attachment to Hohhot with an interesting example.

(5-2) NMF1, New Town, middle-aged group, female

NMF1: I like Hohhot very much. It's not backward, not poor ... We've been living in Hohhot for decades, and we really like here. I remember once we went to Wuhan (a big city in South China), and my husband said, "I will never live here even if you give me \$100,000 a day." So by comparison, we really think Hohhot is good, from the bottom of our heart!

Meanwhile, some speakers were more rational or objective about the positive and negative sides of Hohhot and were more open to the options of moving to bigger cities, thus demonstrating less emotional attachment to Hohhot. In the statistical analysis, this variation in their emotional orientation was found to be a significant predictor for their use of the weak-strong pattern; therefore, it is plausible to say that the middle-aged generation may relate the weak-strong pattern to a sort of "localness" or "Hohhot identity". However, when self-reporting the use of this Jìn feature, the middle-aged speakers rarely said so. Some speakers produced a lot of weak-strong patterns in the elicitation task, but later when I asked them to comment on the stress pattern variation, they often claimed that they never used the weak-strong pattern because they were speaking Pǔtōnghuà, not Jìn. Those who reported their use of the weak-strong pattern often gave the explanation that their own speech was not "perfectly" standard Pǔtōnghuà because they grew up in this dialect-speaking area, claiming that the weak-strong pattern was not a "correct" form and they would use the standard strong-weak pattern or even avoid using l-words (in their words, dialect forms or colloquial forms) when they were speaking formally or with children. For

example, in excerpt (5-3), speaker NMF4 provided some reasons why she sometimes used l-words, but sometimes not.

(5-3) NMF4, New Town, middle-aged group, female

NMF4: (when claiming that she does not use the l-word 提溜 ‘carry something randomly’) No, I don’t use it. For words that are sort of close to dialects, I usually don’t use them, because... I think it’s because of my child (meaning that she does not want her child to learn “non-standard” language).

NMF4: (when claiming that she sometimes uses the l-word 卜拉 ‘repeatedly move horizontally with hand or stick’) Yes, I actually use it. Because, you know, since I was born and grew up in Hohhot, my speech could not be that standard. The local dialect words are sometimes very vivid.

This reflects that the middle-aged speakers are very much struggling with their identities and attitudes. On the one hand, they showed to a certain degree an emotional attachment to Hohhot and considered Hohhot as their hometown because they were born and raised there. As introduced in section 3.2, the migrants in Hohhot were very much separated from their places of origin, so the majority of the second-generation speakers had little knowledge of their parents’ original hometowns. In this sense, they would prefer the label of being an indigenous Hohhotian. However, on the other hand, they also regarded themselves as different from the locally-born residents, thus strongly emphasizing the distinction between the New Town and the Old Town, as well as between Pǔtōnghuà and Jìn dialect. In my fieldwork, I also heard a lot about the conflicts between the two towns from the middle group participants. They were mostly born in the 1950s and 1960s, when both the social conflicts and linguistic distinctions between the New Town and the Old Town were the strongest (more evidence can be found in Jankowiak 2013, Borchigud 1996). Participant NMM4 gave a very vivid example in excerpt (5-4).

(5-4) NMM4, New Town, middle-aged group, male.

XW: When you were little, what was the situation between the New Town and the Old Town? Were there strong social conflicts then?

NMM4: We fought! We fought.. a lot! I remember once some Old Towners came to this side (new town), they stole my cap. So I yelled, “Old Towners took my cap!!” Then all my New Town friends showed up and we had a big fight!

XW: So labels like “New Towner” and “Old Towner” were very strong then?

NMM4: Sure. It’s true.

XW: How could you recognize that they were Old Towners?

NMM4: By dialect! They speak Jìn dialect.

Note that in his description, labels like “Old Towner”, “New Towner” were frequently used, and he even used “this side” to refer to the New Town, emphasizing that the two towns were opposing “sides”. Also, to him, people who speak Jìn dialect would probably be labeled as Old Towners, with whom he would even fight. This unique social and linguistic situation was a crucial part of the environment when the second generation of migrants began to form their own way of speaking. Therefore, it is possible that they would also relate the use of the weak-strong pattern with the label of Old Town. In the interview, middle-aged speakers often constantly emphasized that they were not Jìn speakers and they could only speak Pǔtōnghuà, which could be the reason they claimed not to use the weak-strong patterns. In fact, speakers may adopt the weak-strong patterns to build their Hohhot identity as is shown in the statistical results, but when they were explicitly reporting their use of the weak-strong pattern, they would emphasize their distinction from the Old Town speakers.

In the younger speakers, ATTJIN and ATTOT have significant effects on their linguistic production, but the direction of the effects was opposite to what I had expected. This is

explained by a change of linguistic situation in Hohhot when the younger generation began to form their language. After the 1980s, the promotion of Pǔtōnghuà had influenced the local speakers to a large extent, so all educated urban Jìn dialect speakers could also speak Pǔtōnghuà (Borchigud, 1996). As a result, there was no longer a clear linguistic distinction between Old Town and New Town speakers of the younger generation – they all spoke Hū Pǔ. Even though some of them could still speak Jìn dialect at home, they would switch to Hū Pǔ when talking to their peers. All the younger speakers from New Town in my interviews claimed that they had no Jìn-speaking peers in their childhood. Since language could no longer be used as a salient marker at this time, they never knew whether their friends were originally from the New Town or the Old Town unless they were close enough to know about each other's family background. So although younger speakers are still consciously aware that the weak-strong pattern is from Jìn dialect, they do not directly relate this feature to the local speakers.

From the interviews with the younger generation, it could be seen that the social meaning of the weak-strong pattern for younger speakers might have changed. Many of them claimed that they felt awkward when pronouncing some of the l-words in the standard strong-weak pattern, and they would only do it when speaking to outsiders. Such examples are presented in the excerpts in (5-5) and (5-6) below.

(5-5) NYM1, New Town, younger group, male

XW: Do you say /xu³⁵lu⁰/(s-w) sometimes?

NYM1: No, for this word I won't say/xu³⁵lu⁰/(s-w), unless I'm talking to ... someone from south China, like I was saying it deliberately.

(5-6) NYF7, New Town, younger group, female

NYF7: Yes. I use /tiə⁴³liu⁵⁵/(w-s), or /t^hiə⁴³liu⁵⁵/(w-s).

XW: Do you say /ti⁵⁵liu⁰/(s-w)?

NYF7: Hmm... /ti⁵⁵liu⁰/(s-w)... yes, but very rarely. It's a bit awkward to say that.

...

XW: Do you say /xua³⁵la⁰/(s-w)?

NYF7: Yes, sometimes, but only to outsiders.

Some of the speakers also mentioned why they would prefer to use the weak-strong pattern in some l-words, and they tended to explicitly relate this linguistic feature to their identity of being a local, urban Hohhotian, in order to distinguish themselves from an outsider. In the excerpt in (5-7), Speaker NYF4 mentioned that she would not use the strong-weak pattern in some l-words because it sounded as if she were posturing. And Speaker OYF2¹⁹, in the excerpt in (5-8), playfully accused her friend NYF5 of not being qualified to be a Hohhotian after NYF5 used the strong-weak pattern.

(5-7) Speaker NYF4: younger group, female

XW: Do you use /xua³⁵la⁰/(s-w)?

NYF4: No, it's like you're posturing.

(5-8) Speakers NYF5 and OYF2: younger group, female

NYF5: This rabbit, the ears are (drooping)/ta⁵⁵la⁰/(s-w). That rabbit, the ears are standing upright.

OYF2: As a Hohhotian, you don't say /tə⁴³la⁵⁵/(w-s)?! You're...!

¹⁹ OYF2 is actually an Old Town participant as indicated in the code. She was interviewed together with her friend NYM5 from the New Town. I used OYF2's quotes here because they are very interesting and are explicit comments about using the weak-strong pattern as an identity marker.

NYF5: It's true! I never say /tə⁴³la⁵⁵/(w-s).

OYF2: You are so not qualified (to be a Hohhotian)!

These examples support the claim that for the younger generation speakers in Hohhot, new social meanings might have emerged in the weak-strong pattern of l-words, which was oriented towards their identity of being an urban Hohhotian. As discussed above, this new identity might have already begun to germinate in the middle-aged generation speakers, but they rarely expressed it as explicitly as the younger speakers in the interviews, probably due to their struggling ideologies and deep-rooted prejudice against the Old Town. So this urban Hohhotian identity might be more perceived as a conservative covert prestige for the middle-aged generation. The younger speakers, however, had converted this hidden ideology into an overt prestige and would publicly emphasize their local identity using the weak-strong stress. So there might be a positive correlation between their urban Hohhot identity and their use of the weak-strong pattern. However, why was a significant negative correlation found in their attitudes towards the Old Town (ATTOT) and Jìn dialect (ATTJIN)? This could be due to the fact that speakers who strongly emphasize their urban Hohhot identity tend to care more about what has happened in the city, and are more likely to hear stories and care about the conflicts between the Old Town and the New Town; thus they expressed more negative attitudes towards the Old Town and the local dialect. But they are still more likely to produce the weak-strong patterns to build their urban Hohhot identity. By contrast, for people who do not quite emphasize their Hohhot identity, they may not even care about this social conflict; thus they tend to show more neutral attitudes towards the Old Town community. Of course, to test the changes in the social meaning of the stress pattern variable in the younger generation, perception experiments should be conducted (e.g., Campbell-Kibler 2008 used acoustically manipulated speech recordings to understand the social meaning of individual linguistic variables) and other linguistic variables should be taken into consideration.

5.4.2 Interaction between attitudes and social contact

In the New Town best-fit model, speakers' contact with local Jìn speakers SOCNET was also found to be a significant predictor when interacting with AGE and three attitudinal scores. Figure 5.5 presents the plot of the two-way interaction between SOCNET and AGE. The X-axis is the SOCNET index score, in which larger numbers refer to more contact with Jìn speakers. The Y-axis is still the log odds of probability of producing the weak-strong pattern. The three age groups are represented by the three line types. The plot shows that SOCNET is significantly correlated with speakers' use of the weak-strong pattern in all three age groups. The direction of the effects is the same: speakers having more Jìn speakers in their social networks are significantly more likely to adopt the Jìn feature (older speakers: $p < 0.001$; middle-aged speakers: $p < 0.001$; younger speakers: $p < 0.05$).

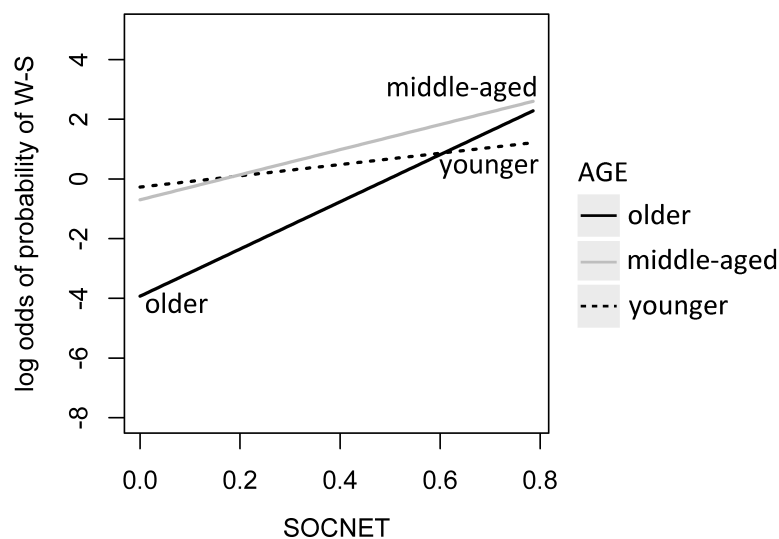


Figure 5.5 Effects of interaction between SOCNET and AGE in the New Town data set.

The best model also found a significant effect of the interaction between SOCNET and speakers' attitudinal scores. This is especially interesting because I would like to see how attitudes and social networks may interact to influence speakers' linguistic production. However, as AGE is not taken into account in this interaction, it is difficult to see the actual effect of these factors. Therefore, to explore these patterns further, the data were further

split into the older, middle-aged, and younger age groups and the effects of the two-way interactions between SOCNET and the attitudes scores were examined separately.

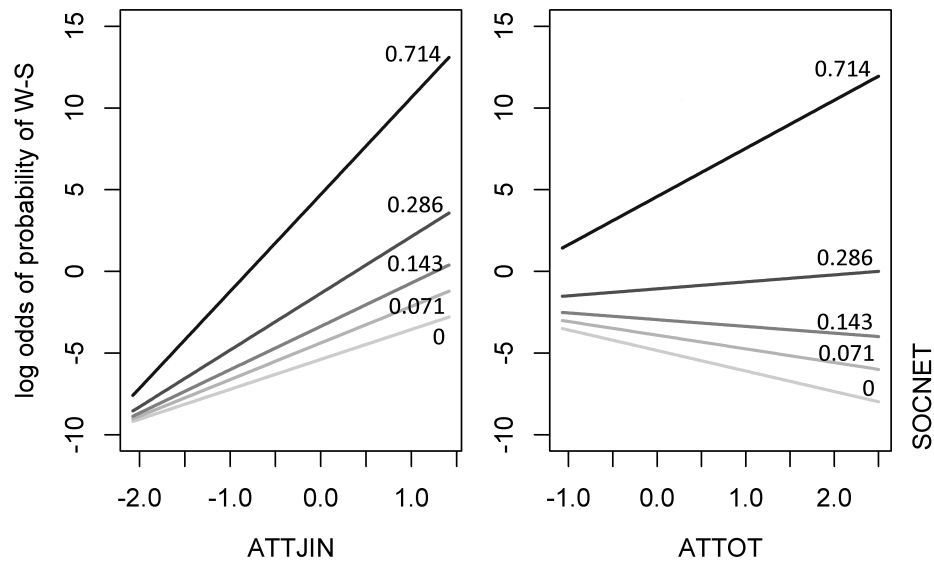


Figure 5.6 Effects of the interactions between SOCNET and attitudinal scores in the New Town older speakers.

In the older speakers, as ATTJIN and ATTOT were found to be significant predictors, the interactions between SOCNET and these two attitudinal variables were examined. Figure 5.6 shows the plots of the effects of these interactions. Again, the Y-axes are the log odds of the probability of speakers using the local *jìn*-feature: a weak-strong stress pattern. The X-axes in the two plots represent the attitudinal variables ATTJIN and ATTOT respectively, where a higher number orients to more positive attitudes. The five grayscale lines show the trend of the effects of SOCNET, in which a higher number (represented by a darker color) means more contact with the local *jìn* community. The five numbers by the lines are the sample minimum, the first quartile, the median, the third quartile, and the sample maximum of the numeric variable SOCNET in the older-group data, which are calculated using the *quantile()* function in R. It can be seen from the plots that: 1) Individuals' social networks influence their stress pattern production. Speakers with more contact with the local community are more likely to use the local weak-strong feature; 2) For speakers who have the same level of

contact with the local community, their attitudes towards Jìn dialect play a role. Speakers with more positive attitudes are more likely to use the weak-strong pattern; 3) The effects of attitudes are stronger in speakers with higher social contact scores because the gradient becomes larger. 4) Speakers with both more positive attitudes and more contact with the local community always have the highest probability of using the weak-strong pattern.

In the previous section when I explained the negative correlation between ATTOT and stress pattern production in the older generation, I attributed the negative direction of the effect to speakers who had very positive attitudes towards the Old Town but had little contact with Jìn speakers. This explanation is now justified by the effect of the interaction between SOCNET and ATTOT (the right plot in Figure 5.6) in the older-generation data. The plot shows that negative correlations are only found in speakers with lower SOCNET scores, because even though they have very positive attitudes towards the Old Town community, they do not have enough contact with them to adopt the weak-strong pattern.

For the middle-aged group, HOHORE was found to have significant effects. The plot of the interaction between SOCNET and HOHORE is shown in Figure 5.7. Again, attitudes are shown on the X-axis, and the effects of social networks are represented by the five trend lines. The five numbers indicate the quantiles (the first, third quartiles, the median and the range) of the SOCNET variable in the middle-aged group. The effects are similar to the results found in the older group: individuals' social contact with local people and their attitudes play independent roles in their stress pattern use. For speakers with the same level of social contact with Jìn speakers, the effect of their future orientation is still found.

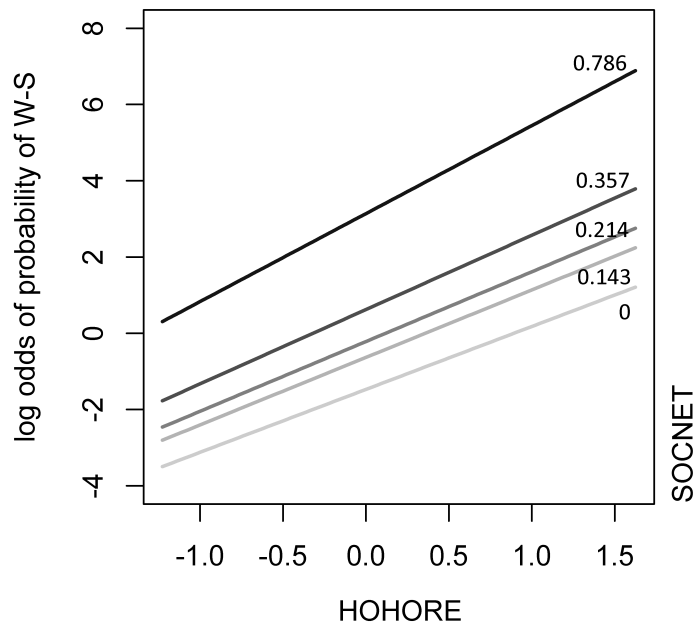


Figure 5.7 Effects of interaction between SOCNET and attitudinal score HOHORE in the New Town middle-aged group.

In the younger group, ATTJIN and ATTOT were found to have significant effects on the stress pattern choice. But the direction of the effects of both factors was negative, which could be related to the change of the social meaning of the weak-strong pattern as explained in the previous section. Plots of the interaction between SOCNET and these two attitudinal scores in the younger generation are shown in Figure 5.8. The social contact effects are still represented by the grayscale lines. Note that there are only three trend lines in this plot rather than five. This is due to the smaller range of the variable SOCNET in the younger speakers. The *quantile()* function still returns five numbers, but the first and third quartiles are the same as the sample minimum and sample maximum, so there are only three lines shown in the plots. The patterns in the plots still demonstrate the effects of individuals' social networks, and the effects of attitudes are retained even when the effect of SOCNET is taken into consideration. Moreover, the general pattern, that speakers with both higher social contact scores and higher attitudes scores are more likely to produce the weak-strong pattern, was attested to.

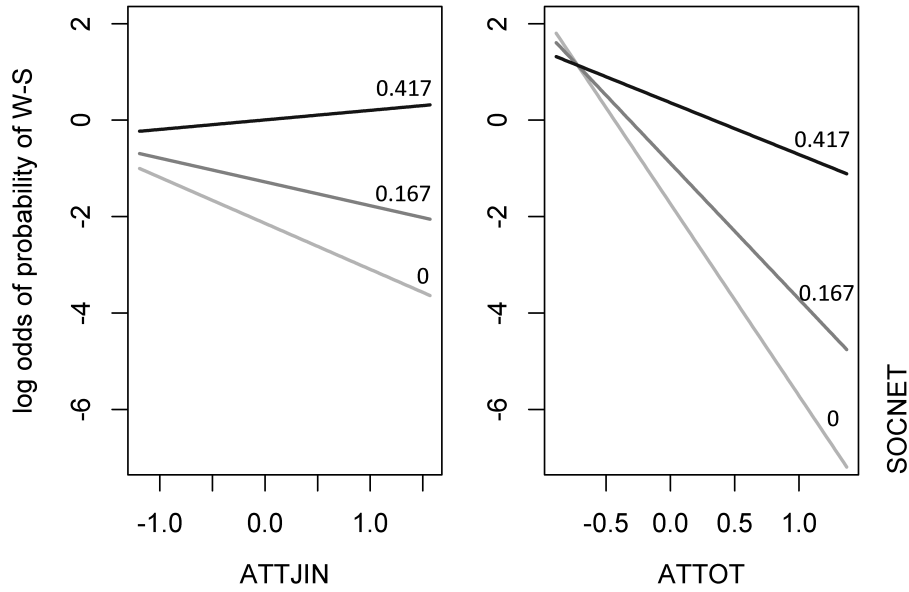


Figure 5.8 Effects of interactions between SOCNET and attitudinal score in the New Town younger generation.

To summarize, the results showed that speakers' social contact with local Jìn speakers is a significant predictor of their adoption of the weak-strong pattern in all three generations. The investigation of the interaction between attitudes and social networks suggests that the effect of speaker attitude may be independent from the effect of social contact, even in larger-scale dialect contact situations like Hohhot. Speakers' social attitudes still play a role even when their social networks are considered. Moreover, the effects of attitudes are often more obvious in speakers with higher social networks scores, and across all speakers and those with both more positive attitudes and more social contact are often more likely to use the local dialect form.

5.5 Results of the Old Town data set

Next, I explore the social constraints operating on speakers' stress pattern use in the Old Town data set. Here, the situation is more complicated in that there are Jìn speakers, Hū Pǔ/Pǔtōnghuà speakers, and bi-dialectal speakers in the Old Town. Table 5.5 shows the

distribution of speakers with different language backgrounds in three age groups. Most informants in the older group are monolingual Jìn speakers. Two speakers who were from the Gōngdà university neighborhood could speak Pǔtōnghuà as well, and both used Pǔtōnghuà in the interview. The middle-aged speakers were mostly bi-dialectal speakers. As introduced in section 5.3, male bi-dialectal speakers in the middle-aged group often spoke Jìn in the interview whereas female speakers tended to speak Pǔtōnghuà or code-switch between Jìn and Pǔtōnghuà. The younger generation all spoke Pǔtōnghuà/Hū Pǔ in the interview, and some of them claimed to use Jìn at home with their parents. The stress pattern in l-words does not vary if the participant is speaking in Jìn dialect, which only produces weak-strong patterns. Therefore, a sub-set of the Old Town data is examined containing only speakers who use Pǔtōnghuà or Hū Pǔ in the interview. This includes non-Jìn speakers and bi-dialectal speakers who did not speak Jìn in the interview. As a result, the sub-set has 610 tokens from 16 speakers, including 2 older speakers, 3 middle-aged speakers and all the 11 younger group speakers.

Table 5.5 Number of Jìn, non-Jìn and bidialectal speakers in each age group.

	Jìn	bidialectal	non-Jìn	total
older	7	2	0	9
middle-aged	3	7	1	11
younger	0	4	7	11

The social predictors tested include AGE, SEX, LANG (whether they were bi-dialectal or non-Jìn speakers), MOVNT (whether they were still in the Old Town or had moved to the New Town), the attitudinal index ATT, and the social networks index SOCNET. Before fitting the models, these predictors were checked for multicollinearity using the VIF scores. LANG and AGE were correlated because older and middle-aged speakers were all bi-dialectal. However, it was not easy to make a decision on whether LANG or AGE should be retained. So in the model fitting process, they were entered into the model alternatively to check which one could better predict linguistic behavior. Other predictors were not highly correlated with each other.

Table 5.6 Output of the best model in the Old Town data set.

	Estimate	Std.Error	p value	
(Intercept)	2.18001	0.68643	0.001494	**
MOVNTno	-1.2917	0.67559	0.055881	.
ATT	1.01848	0.28955	0.000436	***
SEXmale	0.21666	0.50584	0.668417	
SOCNET	-7.11329	1.4892	1.78E-06	***
ATT * MOVNTno	0.04601	0.55611	0.934067	
SEXmale * MOVNTno	-0.97552	1.04302	0.349643	
SEXmale * ATT	-5.19696	1.04672	6.87E-07	***
MOVNTno * SOCNET	7.39144	1.75987	2.67E-05	***
MOVNTno * ATT * SEXmale	6.64512	1.96884	0.000738	***

The 610 tokens were hand fitted into a binomial mixed effects model in R with speakers' stress pattern choice as the dependent variable. The social predictors and the interactions between these predictors were tested as fixed effects. The best-fit model for the Old Town data set had a three-way interaction between ATT, SEX and MOVNT, as well as a two-way interaction between SOCNET and MOVNT. AGE and LANG were found to be non-significant. Speaker and word were included as random intercepts. The social predictors included in the fixed effects were also tested as random slopes for word, but the parsimonious random effects diagnostic technique (Bates et al., 2015; as mentioned in section 4.8) preferred the model with no random slopes included. The output of the best-fit model is shown in Table 5.6, and the plots for the main effects found are presented in Figures 5.9 and 5.10.

The attitudinal score ATT was found to have significant effects on the Old Town speakers' stress pattern use when interacting with SEX and MOVNT. The plots of this three-way interaction are shown in Figure 5.9. The X-axes are the attitudinal score ATT, in which a higher number means more positive attitudes. The Y-axes are still the log odds of probability of the weak-strong pattern. The black lines represent male speakers whereas the gray lines are for female speakers. The left plot shows the interaction of SEX and ATT in speakers who were still living in the Old Town at the time of the interview, and the right plot shows the pattern for speakers who had moved to the New Town. Female speakers

living in the Old Town (the gray line in the left plot) demonstrated a positive correlation between their attitudes and stress pattern production. Speakers who had more positive attitudes towards the local dialect and the Old Town were significantly more likely to produce the weak-strong pattern ($p<0.05$). A similar effect was found in female speakers who had moved to the New Town (the gray line in the right plot), and the effect was even stronger ($p<0.001$). However, the male speakers were behaving in a different way. For speakers who did not move to the New Town (the black line in the left plot), their ATT score was found to be non-significant. For those who were living in the New Town (the black line in the right plot), a negative correlation was found between their ATT score and stress pattern use ($p<0.001$).

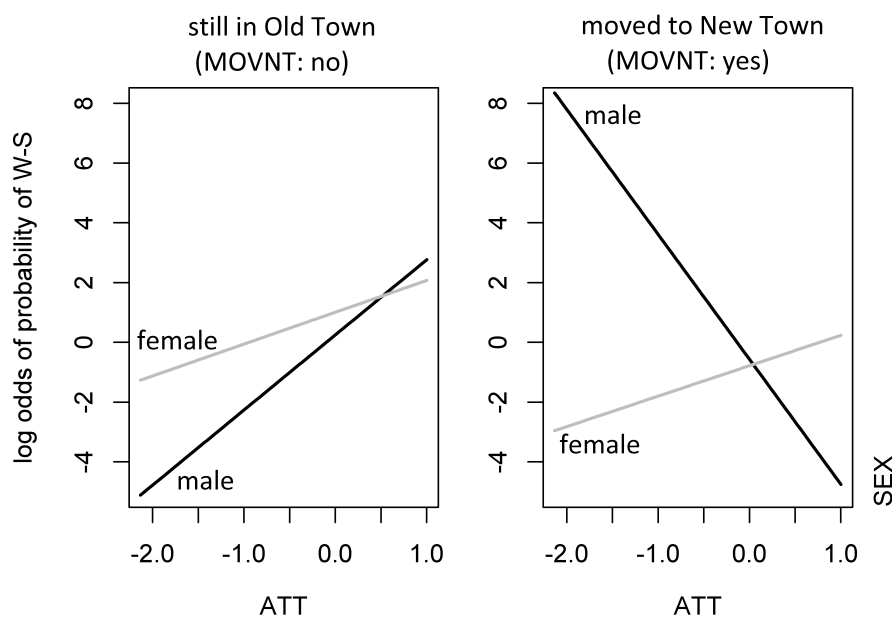


Figure 5.9 Effects of the three-way interaction between ATT, SEX and MOVNT in the Old Town data set.

This difference between male and female speakers could be explained by evidence from the interviews with the Old Town participants. Female speakers who spoke in a more “standard” way with fewer Jìn features in their speech often emphasized the status and “correctness” of Pǔtōnghuà, and insisted that Jìn speakers should learn to speak Pǔtōnghuà; thus they obtained lower ATT scores. For example, speaker OMF4 was a middle-aged female speaker

who was born in the Old Town but later moved to the New Town after being married to her husband whose parents were state-sponsored migrants. In the interview, she constantly emphasized that she had shifted to speak Pǔtōnghuà and had “forgotten” the Jìn pronunciation of words. In the excerpts in (5-9), OMF4 talked about how she shifted from Jìn to Pǔtōnghuà, and why she was reluctant to use Jìn features.

(5-9) OMF4, Old Town, middle-aged group, female

OMF4: I only spoke Jìn when I was very little. Later when I went to high school, the school required us to speak Pǔtōnghuà. I felt awkward at first, but since everybody else was speaking Pǔtōnghuà, I was reluctant to speak Jìn any more.

OMF4: Now I’m living in such an environment, a university, so you can’t use too many Jìn features, or you will sound vulgar... I don’t want other people to (pick on me)...

OMF4: Now that I’ve married him, I have to speak Pǔtōnghuà. I cannot use Jìn dialect any more.

OMF4: (reading the questionnaire) “I’m proud to be an Old Towner.” I disagree!

XW: Do you think you are an Old Towner or a New Towner now?

OMF4: I think I’m a New Towner.

OMF4: (reading the questionnaire) “I’ll be sad if Jìn dialect disappears in the future.” No, I won’t be sad. This dialect (Jìn) is already too unpleasant. ... People should go towards a more civilized and more standard direction.

It can be seen that OMF4 demonstrates very negative attitudes towards the local Jìn dialect. Even though she herself used to be a Jìn speaker, she consciously tried to remove the local dialect feature in her speech, and would prefer to be perceived as a New Towner instead of an Old Towner. The social pressure from her neighborhood, her school, and even her

husband further aggravated her feelings. In the word elicitation task, OMF4 was obviously trying to avoid using local dialect words, or even colloquial words in Pǔtōnghuà. When she produced some weak-strong patterns in l-words, she would immediately emphasize that she only spoke like that when she was little, which may be considered as a sort of “hypercorrection”. A similar case was the speaker OMF2, who also moved to the New Town because of her husband. In excerpt (5-10), she gave an interesting illustration about how she shifted from Jìn to Pǔtōnghuà.

(5-10) OMF2, Old Town, middle-aged group, female

OMF2: I didn’t speak Pǔtōnghuà at all before I got married and moved to the New Town. I learned to speak Pǔtōnghuà because my husband asked me to. When he brought me to dinner with his friends, he asked me not to speak Jìn dialect, because his friends would laugh at him.

This evidence from the interview may explain why female speakers’ stress pattern production is positively correlated with their attitudes towards the local dialect and the Old Town. Speakers who moved to the New Town might be facing even more social pressure to retain Jìn features in their speech. Therefore, if they hold negative attitudes towards the Old Town and Jìn dialect, they may put in more effort to “hide” their accents and identity in their speech, which could sometimes even lead to hypercorrection. This might be why a stronger effect of ATT score was found in speakers having moved to the New Town.

The male speakers, however, presented a quite different pattern. The ATT scores were not significant for participants who did not move to the New Town. Actually in the examined sub-set of the data, there were only three male speakers who still lived in the Old Town and they were all from the younger group. The three of them did not show too much difference in their attitudes or their probability of using the weak-strong pattern, which may be the reason no significant effect was found. For male speakers who had moved to the New Town, the ATT scores were negatively correlated with their stress pattern production. This effect might have resulted from the two older speakers being examined in the sub-set of the Old

Town data: OOM1 and OOM3. Both of them were retired professors from the Gōngdà university, and were both bi-dialectal speakers who chose to speak Pǔtōnghuà in the interview. OOM3 was actually speaking Pǔtōnghuà with a strong Jìn accent, showing a very high probability of using Jìn features in his speech. He was also very active when completing the word elicitation task by giving me lots of examples from Jìn dialect. However, when completing the questionnaires, he expressed very negative attitudes towards the Old Town and Jìn dialect. On the contrary, OOM1 was speaking very standard Pǔtōnghuà with very few local features. He said it could be related to his experience of studying and working in other cities. During the word elicitation task, OOM1 only produced 6 tokens of l-words (all in strong-weak patterns), and insisted that he seldom used these “dialect words”. However, when completing the attitudinal questionnaire, he presented relatively more positive attitudes towards Old Town and Jìn dialect, and obtained a higher score of ATT. This could partly be attributed to a group interview effect. OOM3 was interviewed together with his wife, who demonstrated very negative attitudes towards the local community and dialect. By contrast, OOM1 was interviewed with other two older speakers from the Old Town, who explicitly emphasized their Old Town identity and obtained even higher ATT scores. In both interviews, I was reading out the statements in the AAS questionnaires, and the participants were giving their responses one after another. In some cases they would tend to agree with what others had said, unless they had a very different opinion. Therefore, it is plausible to say that the attitudes of OOM1 and OOM3 might be influenced by other speakers in the same interview.

To sum up, Old Town speakers’ attitudes (ATT) were also found to be a significant predictor for their use of the weak-strong pattern when interacting with SEX and MOVNT. The effects of attitudes in female speakers were as expected, in which speakers with more positive attitudes towards the local dialect and community were more likely to use the weak-strong pattern. However, the male speakers behaved in an unexpected manner and their ATT scores were either non-significant or showed negative correlation with linguistic production.

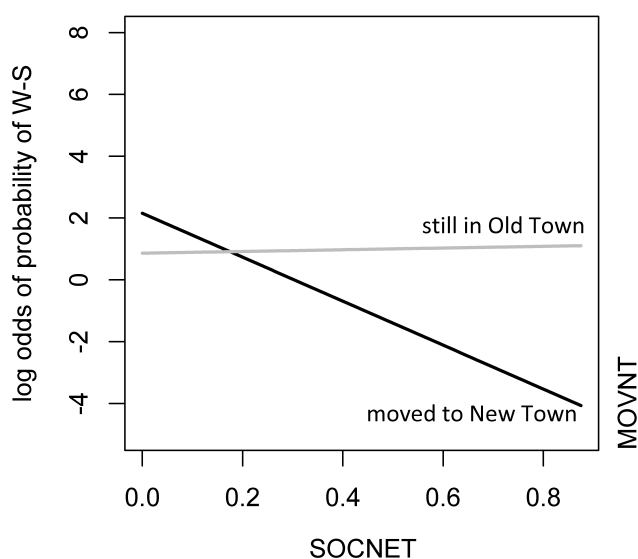


Figure 5.10 Plot of the interaction between SOCNET and MOVNT in the Old Town data set.

In the best model of the Old Town data set, the interaction between MOVNT and speakers' social networks score SOCNET was also found to have a significant effect. The plot of this interaction is shown in Figure 5.10. The SOCNET score is represented by the X-axis, with higher numbers meaning more social contact with Jìn speakers. The black line refers to speakers who have moved to the New Town, whereas the gray line refers to those who are still living in the Old Town.

The result shows that SOCNET is not a significant factor for speakers residing in the Old Town (the gray line). However, for speakers who have moved to New Town (the black line), their SOCNET scores are negatively correlated with their weak-strong pattern use. Speakers who have a higher proportion of Jìn speakers in their networks are significantly less likely to use this dialect feature ($p < 0.001$). As mentioned at the beginning of this section, the speakers included in the current analysis were mostly younger participants. For them, their peer groups at school or work have mostly shifted to speak Pǔtōnghuà, so their SOCNET score actually reflected more about their family background. If their parents were both Jìn speakers, their relatives were often Jìn speakers as well. For the participants who have moved to the New Town, those with Jìn-speaking family background or with higher SOCNET scores may be more likely to be cautious about their speech and consciously alter their

speech to avoid using Jìn features, perhaps to hide their family origin; thus they present a lower probability of using the weak-strong pattern.

5.6 Validity of the attitudinal index scores

In the stress pattern variable results, speakers' attitudinal index scores obtained from AAS were found to be significant predictors of their linguistic production. However, how well can these scores actually reflect speakers' attitudes? This section discusses the validity and reliability of these scores by providing some evidence from the sociolinguistic interview. For the convenience of comparison, I will only focus on a sub-group of participants: the middle-aged speakers from the New Town.

In the New Town results, three attitudinal scores ATTJIN, HOHORE and ATTOT were found to have significant effects. These three attitudinal scores for each speaker in the middle-aged group are shown in Table 5.7. In the table, the highest scores in each attitudinal factor are marked in red, and the lowest scores are marked in blue.

Table 5.7 Attitudinal scores ATTJIN, HOHORE and ATTOT for New Town middle-aged speakers. The highest scores in each factor are marked in red, and the lowest scores are marked in blue.

	ATTJIN	HOHORE	ATTOT
NMF1	-0.57	0.58	-0.26
NMF2	-0.17	1.12	-1.26
NMF3	-0.94	<u>-1.23</u>	-0.64
NMF4	0.53	1.45	-0.20
NMF5	0.40	<u>1.62</u>	-1.08
NMF6	<u>-1.56</u>	0.84	<u>-1.49</u>
NMM1	-0.69	-0.88	<u>0.67</u>
NMM2	-0.37	0.08	-1.04
NMM3	-0.18	0.83	-1.28
NMM4	<u>1.05</u>	1.35	-0.95

For the factor ATTJIN, representing speakers' attitudes towards Jìn dialect, speaker NMM4 obtained the highest score 1.05, and speaker NMF6 had the lowest score -1.56. NMF6 was a librarian at a local university. She was born and grew up in this university neighborhood, and her parents were both state-sponsored migrants originally from Guangdong province. From Table 5.7, we can see that she obtained the lowest score in both the ATTJIN and ATTOT factors. In the interview, she also demonstrated strong negative attitudes towards the Old Town and Jìn dialect. As shown in excerpt (5-11), NMF6 expressed her negative attitudes in a very explicit way.

(5-11) NMF6, New Town, middle-aged group, female

NMF6: When I was little, I didn't like Old Towners at all. I felt like they are mean, petty, selfish, and... anyway, very different from New Towners.

XW: Do you feel like this because you have actual experience with Old Towners, or is this just what you heard from others?

NMF6: I've heard of this before, and later when I had contact with Old Towners myself, I felt the same way. Even today, New Towners still don't like the Old Towners. I feel like we just can't integrate with them. If my daughter is dating an Old Towner, I will absolutely oppose it!

NMF6: I always feel like that side [Old Town] is backward, and many people are illiterates there...

NMF6: Yes, it [Jìn dialect] is vulgar. I think it has a lot of dirty words, and cursing words.

By contrast, speaker NMM4 who had the highest score in ATTJIN showed very positive attitudes towards Jìn dialect. He was running a small business in a local antique market. In the interview, he mentioned that when his customers were Jìn speakers, he would often talk in a very local way or try to speak Jìn dialect with them, so that the customers would

feel closer to him and buy something. In excerpt (5-12), NMM4 showed his interest in Jìn dialect and reported that he would love to imitate it sometimes.

(5-12) NMM4, New Town, middle-aged, male

XW: Do you think Jìn dialect sounds pleasant?

NMM4: Yes! I would like to hear Jìn dialect.

XW: Do you think learning Jìn dialect is interesting?

NMM4: Yes, very interesting. Some words sound very pleasant. ... When I'm communicating with them, I often imitate some Jìn dialect words.

In the attitudinal score HOHORE, which concerns speakers' future orientation or whether they would like to stay in Hohhot, speaker NMF5 earned the highest score 1.62, whereas NMF3 got the lowest score -1.23. As shown in the excerpt in (5-13), NMF5 presented a very strong emotional attachment to Hohhot, by giving evidence from her own experience. She did not choose to leave Hohhot when she married her husband who was from another city, and she even stopped her daughter from leaving Hohhot.

(5-13) NMF5, New Town, middle-aged group, female

XW: Do you want your children to stay in Hohhot as well?

NMF5: Yes, I do hope so. If I had not opposed it, my daughter would have gone to Nanjing (a city in Southeast China) already. It was me who stopped her. ... Her cousin was in Nanjing, but I don't want her to go. I'd like her to stay here. Hohhot is good.

XW: If an outsider is complaining about Hohhot, would you defend it?

NMF5: Yes, I will defend it. If you think we're bad, why do you come here? Go back to where you are! (Laugh) Go to the place you like instead!

XW: If you have an opportunity to live in another city that is similar to Hohhot in economic development, will you choose to go?

NMF5: No, I'm not going. I'd like to stay here. If I wanted, I would have already left Hohhot.

XW: So you actually had opportunities to go?

NMF5: Yes, there were opportunities. When I got married, he (her husband) was in Haibowan (a city in Inner Mongolia), but we did not choose to go there.

Speaker NMF3, who got the lowest score in this factor, demonstrated very different orientations. In excerpt (5-14), she brought up an interesting conversation between her daughter and her to showcase the different feelings they had about Hohhot.

(5-14) NMF3, New Town, middle-aged group, female

XW: If an outsider complains about Hohhot, will you defend it?

NMF3: I might just smile and won't defend it... But my daughter is totally different. Whenever people say something bad about Hohhot, she'll be irritated and defend it for Hohhot. Once we even talked about this. I asked her, "Why do you react so strongly?" And she went, "Why do you NOT react so strongly?" ... So I think that's also why I don't defend it for Hohhot - I don't have that strong emotional attachment.

The third attitudinal score ATTOT was about speakers' attitudes towards the Old Town. As previously mentioned, the middle-aged speakers grew up in the years when social conflicts between the New Town and the Old Town were very intense and evident. Therefore, most speakers in the middle-aged group presented negative attitudes towards the Old Town. Speaker NMF6 who had the lowest scores for both ATTJIN and ATTOT, as discussed above, showed strong negative attitudes towards the locally-born community. Speaker NMM1 obtained the highest score 0.67 in ATTOT. In the interview, he still expressed some

prejudice against the Old Town residents, but at the same time, he also showed some positive attitudes towards the Old Town and felt sorry for the construction in the Old Town, which he thought would destroy the traditional buildings, as shown in excerpt (5-15).

(5-15) NMM1, New Town, middle-aged group, male

NMM1: Old Town used to be very nice. There were many traditional stores along the main street, which showed some traditional flavor. And the food there is much cheaper. Now, the construction in the Old Town has made it not as good.

It can be seen from these cases that speakers' attitudinal index scores were to a large extent in accordance with their psycho-social orientations, and speakers did demonstrate very different social attitudes.

However, there were also cases in which the AAS questionnaires could not fully reflect speakers' attitudes. For example, speaker NMM4, as shown above, obtained the highest score in ATTJIN among the middle-aged speakers, and he himself also demonstrated very positive attitudes towards Jìn dialect in the interview. However, he also emphasized that his attitudes had changed a lot since he came into contact with a lot of Jìn speakers in his business in the antique market. When he was young, he often hung out with a group of childhood friends who all grew up in the same university neighborhood, and they all had very negative attitudes towards the Old Town and Jìn dialect. In excerpt (5-16), speaker NMM4 talked about the change in his attitudes, and also mentioned an interesting story from his group of childhood friends.

(5-16) NMM4, New Town, middle-aged group, male

XW: Do you think Jìn dialect is vulgar?

NMM4: No, I don't think it's vulgar now. But when I was little, I thought it was extremely vulgar! Now, I don't think so.

...

NMM4: When we were little, ... the several of us ... we always hung out together. (talking about their criterion of choosing a girlfriend) First, we don't date Old Towners. Second, we don't date someone from the wool textile factory, because their salary is too low. And third, we don't date someone wearing glasses... (laugh). We were saying that playfully of course, but from this joke, you can see that we had a prejudice against people on that side (Old Town).

In similar cases where a speaker's attitudes had changed in their lifetime, it is difficult for the AAS questionnaires to take into consideration the full attitudinal information. Speaker NMM4 was explicitly aware of this change and reported it in the interview, but other speakers who did not report this were also likely to have similar situations.

Moreover, as previously mentioned, speakers' attitudes collected in the group settings were likely to be influenced by other group members, such as the case of OOM3 and OOM1 mentioned in section 5.5. This was due to the specific situations where the older participants could not complete the questionnaires themselves. In other group interviews where participants were asked to complete the questionnaires separately, their attitudinal scores did not show too much mutual influence.

Therefore, although the attitudinal scores obtained from AAS could not fully reflect speakers' attitudes in some specific cases, they were generally found to be in accordance with speakers' psycho-social orientations in the interview to a large extent. For both New Town and Old Town speakers, their attitudinal scores were found to be significant predictors of their stress pattern use in l-words, suggesting that speakers' consciously offered attitudes are likely to predict linguistic behavior and AAS is a valid and effective technique to collect overt attitudes. The results of the stress pattern variable and the validity of AAS will be further discussed in chapter 7.

6 Fricative variation

In this chapter, I present the second linguistic variable examined in this thesis: the fricative variable. That is, for the plosives [p^h, t^h, k^h] and the glottal fricative [h], speakers in the Hohhot area tend to produce them with a velar fricative [x] following, and this velar fricative could also be realized as a uvular fricative [χ] or a palatal fricative [ç] because of assimilation effects to different places of articulation. This is, essentially, on the plosives, a period of frication. As most of the l-words have plosives [p, p^h, t, t^h, k, k^h] or fricative [h] as initial sounds, this allows me to look at the fricative variable in most of the existing l-words tokens.

As described in section 4.4.1, participants in this study had no explicit awareness about this linguistic feature, and it was never mentioned in the interviews. Therefore, this variable could be very different from the stress pattern feature in terms of the level of awareness and speakers' attention paid to the feature in production. Therefore, exploring speakers' attitude effects in this variable could help us to better understand the interplay between the role of attitudes and awareness in speech production. Also, in the stress pattern analysis, I could not explore the linguistic variation of Jìn speakers because they used the weak-strong pattern exclusively. The fricative variable, as we will see, displays variation in both Jìn and non-Jìn participants, thus allowing me to explore the effects of attitudes in all speakers. This chapter will discuss the fricative variation in detail.

6.1 Description of the variable

The fricative variable has been seldom mentioned in previous literature. Some earlier linguists reported this feature in Jìn-speaking areas, but no study has investigated the feature acoustically or explored its linguistic and social distribution.

Chao (1935) proposed eight types of plosives in Chinese. One of them was mentioned as a type of “combination sound” like [p^x]²⁰, which combines a plosive [p^h, t^h, k^h] with a velar fricative sound, as in [p^xa⁵¹] (怕, afraid). He claimed that this type of plosive could be found in many dialects in Shanxi province (the major Jìn dialect speaking area), and the velar fricative would become a palatal fricative [ç] when it is followed by a high front vowel [i], such as in [p^çi³⁵] (皮, skin). He then further explained that, in this type of plosive, aspiration was the major function, and the insertion of fricatives [x] or [ç] was random. Actually, this type of combination aspiration of [p^x, t^x, k^x] was reported long ago by Karlgren (1915-28), who was regarded as the first person in modern linguistics to have investigated Jìn dialects (Qiao, 2005). He claimed that this type of plosive could be found in many Jìn-speaking areas, such as Shanxi, Shaanxi, and Guihua 归化 (the name for Hohhot Old Town at that time) (p. 173), which could be early and direct evidence to prove the existence of this feature in the Hohhot area. More recently, Hou (1999) listed the feature of pronouncing [p^h, t^h, k^h] followed by a voiceless velar fricative [x] as one of the key features of Jìn dialects.

Interestingly, Karlgren (1915-28) mentioned that some other linguists had claimed that this type of plosives could be found in Beijing as well, but he insisted that the standard Beijing pronunciation of plosives was “pure” aspiration without the velar fricative [x], because, according to his experience, well-educated Beijingers always claimed that the Beijing dialect sounded quite “soft” while outsiders’ pronunciation was “stronger”. Karlgren

²⁰ This sound is written as [px] in Chao 1935. Similarly, the following two examples are written as [pxa⁵¹] and [pçi³⁵].

pointed out that what they meant by “soft” and “strong” especially referred to the quality of [p^h][t^h] and [k^h]. However, this discussion suggests that early in Karlgren’s time, the feature of combining [p^h, t^h, k^h] with the fricative [x] could also be heard in non-Jìn areas like Beijing, but maybe not as often as, or as noticeably as, in Jìn-speaking areas. Further evidence of this could be found in Duanmu (2007). When he was describing the inventory of consonants in Pǔtōnghuà, Duanmu mentioned that the aspirated consonants [p^h, t^h, k^h, ts^h, tɕ^h] were often [p^x, t^x, k^x, ts^x, tɕ^x] before a back vowel. By saying that, he might be simply referring to the co-articulation effect where [h] assimilates to the following back vowel and becomes the velar fricative [x]. But this may also imply that a period of frication could be involved in this situation, which would support the claim that this feature is also likely to be found in Pǔtōnghuà, and furthermore, it might be influenced by the place of articulation of the following vowel.

However, all the descriptions about this feature have traditionally been only based on these linguists’ personal observations. Also, most of the discussion was about the plosives [p^h, t^h, k^h], but the phoneme [h] could also have a similar variation.

The [h] phoneme in Pǔtōnghuà, as in hǎo (好, good), is usually phonetically transcribed as the voiceless velar fricative [x], which may indicate that this phoneme in Pǔtōnghuà is different from the English glottal fricative [h], but more similar to the velar fricative [x] “ch” sound in e.g. German. However, Duanmu (2007) has argued that the [x] phoneme in Pǔtōnghuà is also variable, since [x] is sometimes considered to be [h] according to the place of articulation of the following vowel.

This suggests that for both the [p^h, t^h, k^h] phonemes and the [h] phoneme, there could be variation in terms of whether a velar fricative sound [x] is present in the pronunciation. This velar fricative is also likely to be realized as a uvular fricative [χ] or a palatal fricative

[ç] due to the linguistic environment²¹. This feature is noticed by some linguists, perhaps because of the acoustic differences between [h] and [ç, x, χ]. Shadel (1991) and Johnson (2002) make a distinction between “obstacle fricatives” and “non-obstacle fricatives”, where obstacle fricatives all involve some processes in which a jet of air hits an obstacle in the vocal tract, for example, wall source fricatives like [x], obstacle source fricatives like [s], and lip source fricatives like [f, v]. Johnson (2002) states that the only non-obstacle fricatives are bilabial and glottal fricatives. Sounds at these positions are difficult to be made loudly because the airstream does not hit any obstacles. Therefore, the glottal fricative [h], as a non-obstacle fricative, could have very different articulation processes and perceptual effects, compared to the obstacle fricatives [ç, x, χ], which perhaps makes phonetic variation more noticeable. I return to this later, since it is part of the reason I treat this feature as a binary variable, later in the analysis.

There are a number of questions that could be further explored in relation to this fricative variable. Is it simply as what Chao (1935) has described as random variation? Or is it a co-articulation effect proposed by Duanmu (2007) which only relates to the place of articulation of the following vowel? Or will other linguistic factors also influence the realization of this variable, like the stress pattern in l-words? And what about social factors?

Before exploring these questions, there are two issues I need to make clear about this linguistic variable.

²¹ Chao (1935) and Karlgren (1915-28) both mentioned that the fricative [x] would be realized as [ç] when it is followed by a high front vowel [i]. It could also be seen from my data that some tokens were realized as [ç] or [χ], but how this variation is constrained by the linguistic environment needs further investigation.

First, the fricative token [h] and the plosives tokens [p^h, t^h, k^h] are analyzed in the same way in this thesis. As has been described earlier in this section, the variation of [h] and [p^h, t^h, k^h] are discussed separately in previous literature. While [p^h, t^h, k^h] followed by [x] has been recognized as a Jìn feature, the variation between [h] and [x] is seldom mentioned and only described by Duanmu (2007) as a Pǔtōnghuà feature. Also, if we look at the variation of [h] and [p^h, t^h, k^h] in theories of phonological lenition, they seem to be different processes. The classic lenition process, as in the “strength of hierarchies” proposed by Lass (1984, p. 178), for example, is from voiceless stop to affricate to oral fricative then to glottal fricative, as shown in (6-1).

(6-1) voiceless stop -> aspirate/affricate -> oral fricative -> glottal fricative -> ∅

Therefore, the realization of [p^h, t^h, k^h] as [p^x, t^x, k^x] could be seen as a lenition process from voiceless stop to affricate. On the contrary, the realization of [h] as [x] is more of a strengthening process towards the opposite direction: from glottal fricative to oral fricative. So for these reasons, it is possible that the variation of [h] and [p^h, t^h, k^h] will have different linguistic and social distributions. Therefore, in this thesis, all the statistical analysis has been done on the data sets of [h] tokens, [p^h, t^h, k^h] tokens, and the combined data set respectively. However, the results showed no difference in the three data sets, indicating that these phonemes were linguistically and socially behaving in a similar way in my data. Thus, I do not make a distinction between the [h] and [p^h, t^h, k^h] phonemes, and as we will see in the following sections, I will only report the results when the two data sets are combined as one.

Second, for the convenience of describing this linguistic variable, in the following sections, I will only use the velar fricative [x] to represent the obstacle fricatives involved in this feature, but remember that for each token, the [x] sound could also be [ç] or [χ] due to possible coarticulation effects.

6.2 Methodology

Twenty l-words with $[p^h, t^h, k^h, h]$ initial sounds were selected as target words. Altogether 1957 tokens were labeled and analyzed in Praat.

For the plosives $[p^h, t^h, k^h]$ tokens, the duration from the burst of the plosive to the end of the fricative was labeled as an interval (see an example of $[p^h]$ in Figure 6.1). And for $[h]$ tokens, the duration of the entire fricative was labeled (see Figure 6.2).

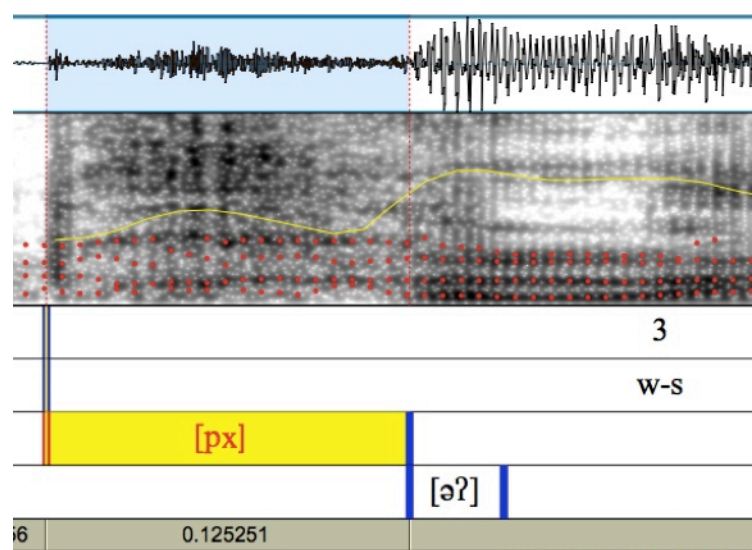


Figure 6.1 Example of the labeling of plosive $[p^h]$ tokens.

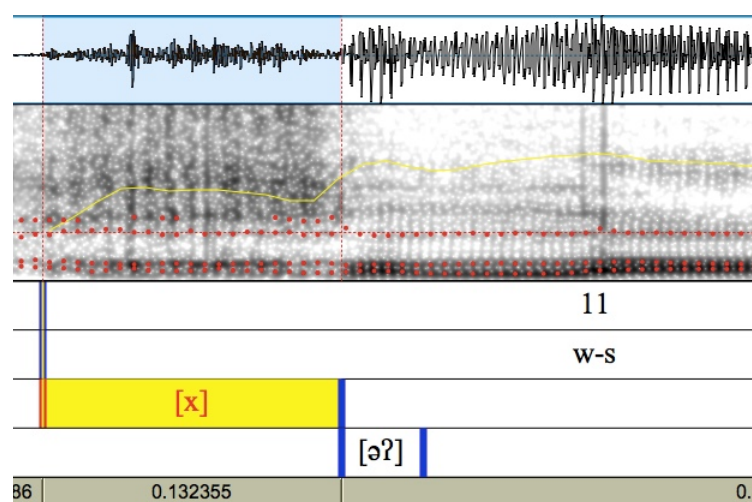


Figure 6.2 Example of the labeling of fricative $[h]$ tokens.

For the coding of these tokens, an impressionistic method was first applied. By auditory analysis, I labeled each token as whether or not a fricative [x] was involved in the production, which formed a binary linguistic variable. However, perceptual coding of the fricative variable was more difficult than the stress pattern variable. Many of the tokens were not straight-forward in terms of whether I could hear a [x] or not. Therefore, some acoustic measurements were made in Praat in order to justify my perceptual coding.

The following six acoustic measures, which have been used in many previous studies of fricatives (e.g., Wright, Hargus, & Miller, 2008; Gordon, Barthmaier, & Sands, 2002), were taken for each token: center of gravity (hereafter, COG), standard deviation (hereafter, SD), skewness, kurtosis, normalized intensity, and duration. All measures were made over the middle 60% of the labeled interval of each token, except for duration, which was measured across the entire interval. I further explain these acoustic measures below.

The measure of intensity refers to how much energy is in the fricative. As previously mentioned, the non-obstacle fricative produced at the glottis, i.e. fricative [h], is very difficult to produce with a high amplitude. If a token has an obstacle fricative [x] in pronunciation, the intensity measures are likely to be higher. After intensity measures were extracted using Praat, they were then normalized by speaker using the *scale()* function in R. This was to eliminate some individual effects, like the distance of the microphone, speakers' loudness, and different interview environments. The duration of each token was also measured. Though it was hard to make a clear prediction about whether [p^x, t^x, k^x, x] and [p^h, t^h, k^h, h] tokens would have different durations, I expected the tokens with [x] to be longer, because there might be more acoustic energy involved when obstacle fricatives were produced and the turbulent airflow was escaping through a narrower channel (Johnson, 2002).

The measurement of COG and SD are related: COG is the average frequency of the entire spectrum, while SD gives us information about how far each frequency value is from that average. The differences could usually be seen visually from the spectrograms. For example,

Figure 6.3 shows the spectrograms of [h] and [x] produced by the same speaker NMF1. The frequency range on the vertical axis is larger in the [x] sound, and the spectral energy is concentrated at higher frequencies for [x] than for [h]. Therefore, it will be reasonable to find tokens with [x] to have higher COG and SD values.

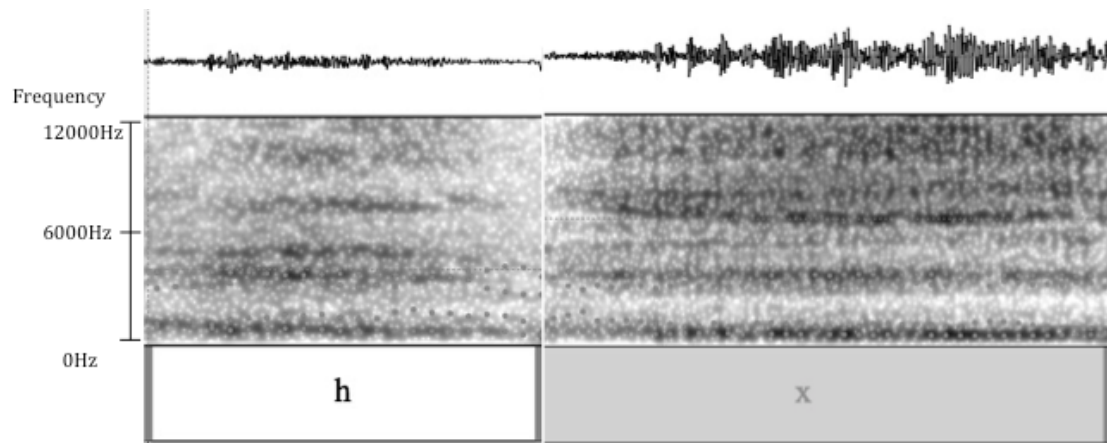


Figure 6.3 Spectrograms of [h] and [x] produced by the same speaker NMF1.

The measure of skewness and kurtosis are both related to the shape of the acoustic spectra of fricatives. Skewness is a measure for how much the shape of the spectrum below the center of gravity is different from the shape above the mean frequency, while kurtosis is a measure for how much the shape of the spectrum around the center of gravity is different from a Gaussian shape. This could be detected from the spectral shapes extracted from Praat. For instance, Figure 6.4 shows the spectral slices across the 60% window of the two tokens [x] and [h] presented in Figure 6.3, which were still produced by the same speaker NMF1. By comparing the two spectra, we can see that for the [h] token shown on the right graph, the amplitudes on the higher frequencies around 10kHz obviously drop, which makes the shape more skewed and more different from a Gaussian shape. Thus, I predicted that tokens with a velar fricative [x] will have lower skewness and kurtosis values than those without [x].

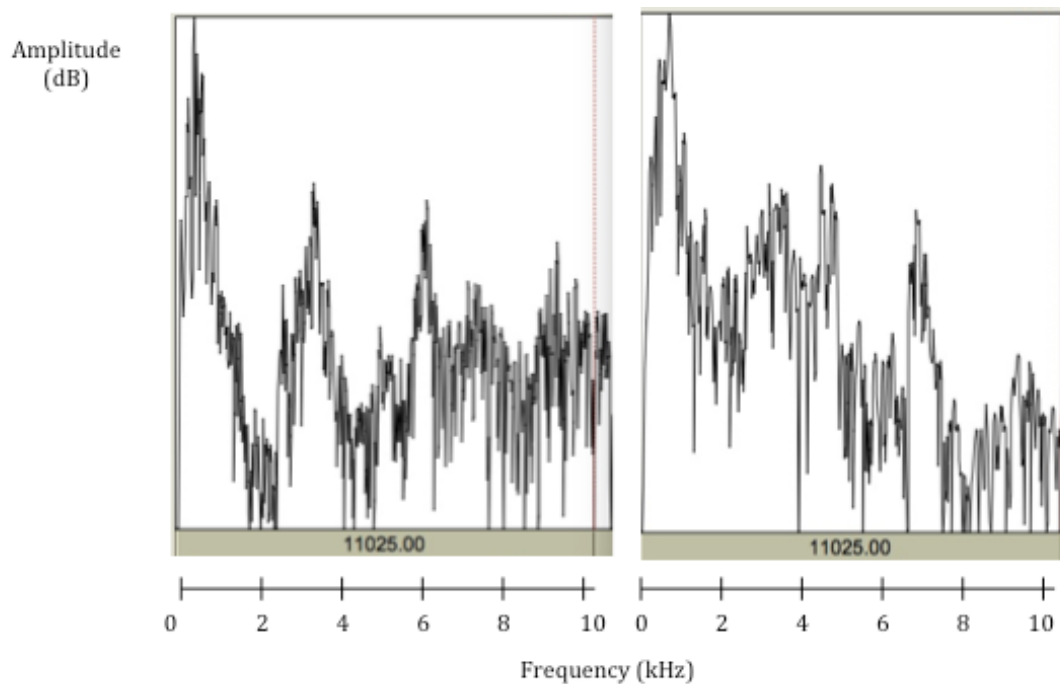


Figure 6.4 Acoustic spectra of the same two tokens [x] (left) and [h] (right) produced by speaker NMF1.

After obtaining the six acoustic measures for each token, they were used to corroborate my perceptual coding. This step was done by fitting regression models to check whether these acoustic measures are significantly correlated with my auditory coding. However, it should be clear from the discussion above that some of the measurement values are likely to correlate with each other. In order to avoid multicollinearity and reduce the dimensions of the six measures, a principal component analysis (PCA) was run on these measures, again, using the *principal()* function in the *psych* package in R. The result revealed four main factors (see Table 6.1). The first factor PC1 is related to the measure of skewness and kurtosis. PC2 and PC3 are related to normalized intensity and duration respectively. The last factor PC4 is relevant to COG and SD values. Four PC scores were then calculated for each token to represent their acoustic values in these four aspects.

Table 6.1 Four acoustic measurement factors revealed by PCA.

Factors	Representing acoustic features
PC1	skewness and kurtosis
PC2	normalized intensity
PC3	duration
PC4	COG and SD

Then, a binomial mixed-effects model was fit in R with the perceptual coding results as the dependent variable. The four PC scores revealed by PCA representing different dimensions of the acoustic features were tested as independent variables, as well as two linguistic variables PHONEME (the phonemes [p^h, t^h, k^h or h]) and FVOWEL (the following vowel). The outcome of the best-fit model is shown in Table 6.2, in which the four PC scores are all significantly correlated with the auditory coding. The results showed that tokens are significantly more likely to be perceptually coded as “with [x]” if they had higher COG and SD values ($p < 0.001$), higher intensity ($p < 0.001$), longer duration ($p < 0.001$), and lower skewness and kurtosis values ($p < 0.01$), which was exactly as predicted in the earlier analysis. The result indicates that my auditory coding was justified and could be used for further analysis.

Table 6.2 Outcome of the best model for correlation between auditory coding and acoustic measures.

	Estimate	Std.Error	p-value	
(Intercept)	-3.21862	0.59559	6.52e-08	***
FVOWELi	-8.35718	2.60057	0.00131	***
FVOWELə	1.24373	0.39919	0.00184	***
FVOWELu	1.23624	0.41360	0.00280	***
PHONEMEH	1.93128	0.45891	2.57e-05	***
PHONEMEp	0.61943	0.53002	0.24253	
PHONEMEt	1.23901	0.50765	0.01466	*
acoustic_pc1	-0.27315	0.09209	0.00302	**
acoustic_pc2	1.34279	0.09754	< 2e-16	***
acoustic_pc3	0.65606	0.09066	4.61e-13	***
acoustic_pc4	0.60071	0.10413	7.99e-09	***

As the auditory coding is verified, the social and linguistic constraints operating on the fricative variable are then explored by using the binary perceptual coding data. In this variable, I am still interested in the potential effects of speakers' attitudes and social networks, and how the attitudes effects might be different for different generations of speakers in the koinéization process. And particularly, as the fricative variable is a feature that is below speakers' conscious awareness, the results of this feature will contribute to the research question of whether explicit awareness is a threshold for attitudes to play a role in language change (research question 7).

As with the stress pattern variable, the effects of attitudes on the fricative variable could only be explored in separate data sets of New Town speakers and Old Town speakers, because the attitudinal scores were generated from different questionnaires. Therefore, the fricative variable was also first investigated in the combined data set, and then in the New Town and Old Town data set separately. The results are presented in sections 6.3, 6.4 and 6.5.

6.3 Results of the full data set

In the combined data set analysis, all the 1957 tokens produced by both the New Town and the Old Town speakers were hand fitted into binomial mixed effects models in R. The dependent variable was the $[p^h, t^h, k^h, h]$ phonemes being realized as $[p^x, t^x, k^x, x]$, that is, with a period of frication. The independent variables included both social and linguistic factors. For linguistic factors, the following vowel (FVOWEL) was included to test the coarticulation effects proposed in previous literature (Duanmu, 2007), and the phoneme itself was also included as a categorical factor PHONEME to see whether $[p^h]$, $[t^h]$, $[k^h]$ and $[h]$ tokens would behave differently. The stress pattern of each token (STRESS) was also tested as a potential predictor. As for social factors, the effects of AGE, SEX, TOWN, EDU, OCCU and RELATION were tested.

Before the modeling, the VIF statistical diagnostic method was used again to check multicollinearity. The result showed that the two linguistic factors PHONEME and FVOWEL were weakly correlated with each other. This is due to the fact that, in the selected l-words, some vowels would only appear with some specific phonemes. For example, the [i] vowel could only be found in tokens like [t^hiəʔ⁴³liu³¹] “carry something randomly”, so it could only appear with the [t^h] phoneme. The [a] vowel mostly appeared with the phoneme [t^h] as well in l-words like [t^ha⁵⁵la⁰] ‘wear shoes like slippers’. This sort of collinearity is related to the words themselves, which is hard to avoid. So here, both factors were kept in the model, but when explaining the results, it is important to note that the individual effects of FVOWEL and PHONEME are hard to detect here because they are originally correlated in the selected words. To further test these effects more accurately, the [p^h, t^h, k^h, h] phonemes should be examined in a broader linguistic environment, as in spontaneous speech instead of just l-words, which is not the focus of the current study. RELATION was also removed because it was highly correlated with AGE. The other independent variables all obtained a VIF value below 3.

The model was again hand-fit using forward stepwise selection, with pairwise model selection guided by the Akaike Information Criterion. For random effects, speaker and word were still included as random intercepts. However, in the model fitting process, it was found that the word intercept was always close to zero; and when it was removed, the models became significantly better. This could have resulted from the fact that FVOWEL, PHONEME and STRESS were included in the fixed effects, which could have explained most of the variation between different l-words selected. Thus the inclusion of word as random intercept might be redundant and made the model worse. For example, the l-words [xuəʔ⁴³lu³⁵] 葫芦 ‘a calabash’ and [xuəʔ⁴³la³⁵] 划拉 ‘scribble’ shared the same phoneme [h] and following vowel [u] in the first syllable and the same stress pattern. The model result suggests that these two words may behave in the same way in terms of the fricative variation, even though their second syllables were different. That is to say, the fricative variation was not related to individual word effects – for any l-words, as long as they share

the same first syllable, they would behave similarly in the fricative variation. Therefore, for all models in the fricative variable analysis, the word intercept was removed and only speaker was included as random intercept.

Table 6.3 Output of the best model for fricative variation in the full data set.

	Estimate	Std.Error	P-value	
(Intercept)	-5.1316	1.2875	6.73E-05	***
PHONEMEH	2.3684	0.5421	1.25E-05	***
PHONEMEp	-1.1641	1.2182	0.339269	
PHONEMEt	1.1186	0.6021	0.063167	.
STRESSw-s	0.6794	0.5064	0.179753	
FVOWELə	1.3858	0.6461	0.031962	*
FVOWELi	-4.9442	2.3174	0.032886	*
FVOWELu	0.2357	0.6365	0.711095	
SEXmale	1.7479	1.0161	0.0854	.
TOWNoldtown	4.687	0.9919	2.30E-06	***
AGEmiddle	2.9927	0.8980	0.00086	***
AGEyounger	2.2657	0.9407	0.016019	*
PHONEMEH * STRESSw-s	-1.3622	0.5475	0.012848	*
PHONEMEp * STRESSw-s	0.7113	1.2258	0.561737	
PHONEMEt * STRESSw-s	-0.1631	0.6289	0.795384	
SEXmale*TOWNoldtown*AGEmiddle	4.6378	1.4971	0.00195	**
SEXmale*TOWNoldtown*AGEyounger	4.4894	1.6844	0.007692	**

The best model had the fixed effects of FVOWEL, a two-way interaction between PHONEME and STRESS, and a three-way interaction between TOWN, SEX and AGE. For random effects, the model has speaker as a random intercept and a by-speaker random slope for PHONEME and FVOWEL. The output of the best-fit model is presented in Table 6.3. Plots of the effects of the main factors are shown in Figures 6.5 and 6.6. The Y-axes in these plots all represent the log odds of the probability of [p^h, t^h, k^h, h] being realized as [p^x, t^x, k^x, x]. The larger the number, the more likely that the fricative [x] is pronounced in the production. The left plot in Figure 6.5 illustrates the effects of the interaction between PHONEME and STRESS. The phonemes [p^h, t^h, k^h] did not show a significant difference in the weak-strong and strong-weak stress patterns in terms of whether a [x] sound is involved. But for the [h] phoneme, it is more likely to be pronounced as [x] in the strong syllables (p<0.05), that is, in the case of a

strong-weak stress pattern. This is easily explained from the perspective of articulation. It could be more likely that the [h] phoneme is realized as [x], which is more intense and “noisy” acoustically, in a stressed and longer syllable. The effect of FVOWEL is shown in the right plot in Figure 6.5. As previously mentioned, Duanmu (2007) proposed that the realization of [x] could be related to the following back vowels, like [u]. From the current result, [u] significantly favors the [x] variants compared to the front vowel [i] ($p < 0.05$), which partly supported Duanmu’s argument. But the vowel [a] is behaving similarly with [u], and [ə] even favors [x] significantly more than [u] ($p < 0.001$). This indicates that the fricative variation is correlated with the following vowel, and it is not simply a coarticulation effect. However, since there is multicollinearity between the following vowel and phoneme, the effect might also be influenced by the corresponding phonemes. Therefore, I cannot make a very strong claim here, but at least it is clear that the following vowel is a significant predictor for the fricative variable.

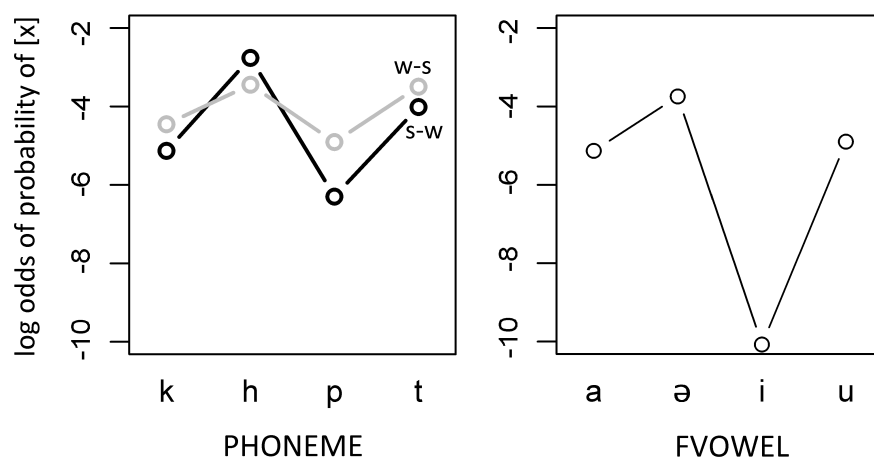


Figure 6.5 Plots of the effects of FVOWEL and the interaction between PHONEME and STRESS for the fricative variable in the full data set.

As for the social factors, the plots for the three-way interaction between TOWN, SEX and AGE are shown in Figure 6.6. The X-axes are the three age groups or generations so that we could see the change in this variable over time clearly. The black lines refer to the New

Town speakers' results, and the gray lines are the Old Town speakers. The two plots show the results from female and male speakers respectively.



Figure 6.6 Plots of the 3-way interaction between TOWN, SEX and AGE for the fricative variable in the full data set.

For the female speakers (the left plot in Figure 6.6), the pattern of the fricative variable shows clearly the typical koinéization process, and the focusing seems to happen in the second generation. The older generation of local Old Town residents, who are mostly Jìn speakers, were significantly more likely to produce the variants with [x], whereas the older speakers from the migrant community were significantly less likely to produce this feature ($p < 0.001$). This result also provides direct evidence for the frequent use of $[p^x, t^x, k^x, x]$ as a Jìn feature, which supports the previous literature mentioned in section 6.1. However, when it comes to the second and third generations, speakers from the Old Town and the New Town seem to have converged to each other and show a similar likelihood of using the [x] variants. This pattern is very similar to the stress pattern variable results presented earlier in section 5.3.

The male speakers (the right plot in Figure 6.6) show a different pattern in the fricative variable. Speakers of the New Town and the Old Town do not show a significant difference,

even in the older group speakers. The reason may still relate to the participants' distribution in the older group. As discussed in the stress pattern results (see section 5.3), the five male participants from the Old Town are actually retired professors in the Gōngdà neighborhood, who cannot be considered as representative speakers of the Old Town older group. Their pronunciation may have converged to non-Jìn speakers consciously or unconsciously due to their social networks and higher educational level. At the same time, the male older speakers from the New Town, who demonstrated more positive attitudes and a higher degree of integration into the local community compared to older females, also presented a higher probability of adopting Jìn features. This trend was also found in the stress pattern variation results – male older speakers from New Town and Old Town tended to converge to each other. However, the stress pattern results still found a significant difference between New Town and Old Town male older speakers, because Jìn speakers' linguistic change could not be revealed in the stress pattern variable – they would use the weak-strong pattern exclusively if they were speaking Jìn dialect. Nonetheless, when it comes to the fricative variable, Jìn speakers are also likely to vary in their production. Therefore, in the fricative variable results, no significant difference was found between New Town and Old Town male speakers, indicating that speakers from both towns were converging to each other in the use of the [x] variable.

This pattern continued in the middle-aged male speakers, in which New Town and Old Town speakers showed a similar probability of using the [x] variants. However, the younger speakers presented some changing trends. For the New Town speakers, the younger group was significantly less likely to use the Jìn form compared to the middle-aged group ($p < 0.05$). From the pattern, it seems that the New Town younger group speakers are leading the change to reduce the use of the [x] variants, and the Old Town younger group speakers are conforming to this trend.

To sum up, the fricative variation results in the full data set show that this variable is constrained by both linguistic and social factors, indicating that the fricative variable is not simply a haphazard or random insertion of [x] as suggested by Chao (1935), and it is also not

a simple co-articulation effect proposed by Duanmu (2007). The language change pattern of this variable was found to be different in male and female speakers, which corresponded to the stress pattern results. The female speakers presented a typical koinéization process in which the speakers from the New Town and the Old Town converged in the second generation, whereas the male speakers did not show a significant difference between speakers from the two towns, which could be explained by the participants' situation in my fieldwork. Additionally, the pattern of the older group female speakers further proves the claim that the frequent use of [p^x, t^x, k^x, x] is a Jìn feature. In the next two sections, I will explore whether speakers' attitudes and social networks will influence their production of the fricative variable in the New Town and Old Town data set separately.

6.4 Results of the New Town data set

In the New Town data set, 1010 tokens from 35 speakers were analyzed. Apart from the social and linguistic predictors examined in the full data set model, speakers' four attitudinal index scores ATTJIN (attitudes to Jìn dialect), HOHORE (future orientation), ATTOT (attitudes to the Old Town), IDMIG (identity of state-sponsored migrants) and their social interaction score SOCNET were also tested as fixed effects. Similarly to the stress pattern variable analysis, the predictors IDMIG, OCCU, RELATION were removed because of multicollinearity checked by the VIF diagnostic method. The best model has the fixed effects of the following vowel FVOWEL, PHONEME, speakers' social networks score SOCNET, and two-way interactions between AGE and three attitudinal scores ATTJIN, HOHORE and ATTOT respectively. Speaker was included as the random intercept, and word intercept was removed because of the same reason described in the previous section. The outcome of the best model is presented in Table 6.4, and the plots of the main effects are shown in Figure 6.7 to 6.9.

Table 6.4 Output of the best model for the fricative variation in the New Town data set.

	Estimate	Std.Error	p-value	
(Intercept)	-3.09416	0.51332	1.66E-09	***
FVOWEL _a	0.90985	0.30998	0.003333	**
FVOWEL _i	-2.81448	0.78002	0.000308	***
FVOWEL _u	-0.01247	0.3217	0.969073	
PHONEME _h	1.27151	0.30831	3.72E-05	***
PHONEME _p	-0.53803	0.39711	0.175462	
PHONEME _t	0.76775	0.35478	0.030464	*
ATTJIN	0.30415	0.22472	0.175905	
HOHORE	-0.47479	0.28476	0.095452	.
ATTOT	0.404	0.26629	0.129231	
AGEmiddle	1.44203	0.32897	1.17E-05	***
AGEyounger	2.0147	0.40562	6.80E-07	***
SOCNET	1.70652	0.54416	0.001712	**
AGEmiddle * ATTJIN	-0.98847	0.38261	0.009781	**
AGEyounger * ATTJIN	-0.99453	0.38586	0.009954	**
AGEmiddle * HOHORE	0.68082	0.36789	0.064223	.
AGEyounger * HOHORE	1.91608	0.4106	3.06E-06	***
AGEmiddle * ATTOT	0.08092	0.37513	0.829214	
AGEyounger * ATTOT	-3.1708	0.54918	7.75E-09	***

Figure 6.7 presents the effects of the linguistic predictors: FVOWEL and PHONEME. The pattern of the following vowel is almost the same as that in the full data set. Phoneme also displays a similar effect, but unlike in the full data set, it is not interacting with stress pattern. Actually, stress pattern was found to be non significant in the New Town models. As for the effect of phoneme, the [h] phoneme is significantly more likely to be realized as [x] compared to [p^h] (p<0.001), [t^h] (p<0.05), and [k^h] (p<0.001) respectively. The phoneme [t^h] is also more likely to be pronounced with [x] than [k^h] (p<0.01) and [p^h] (p<0.001). This further proves that the fricative variation is not simply a coarticulation effect.

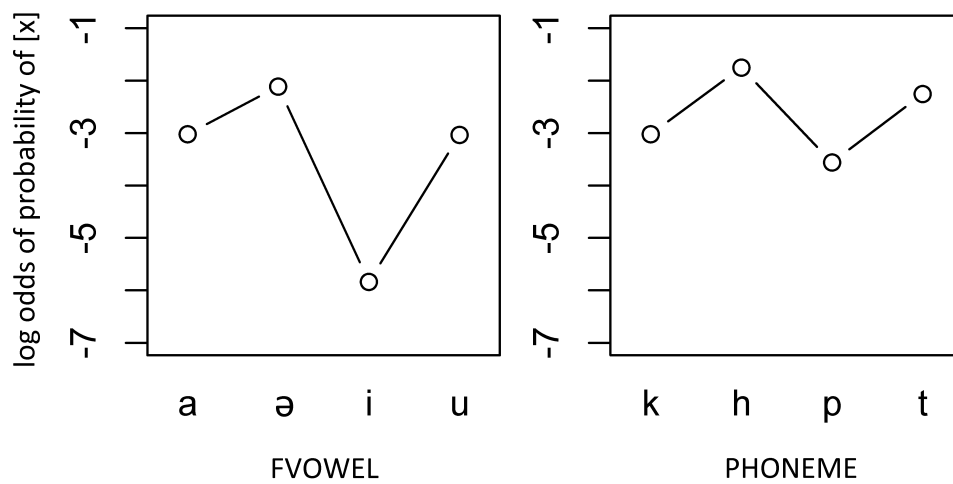


Figure 6.7 Plots of the effects of FVOWEL and PHONEME in the New Town data set.

For the social predictors, speakers' social interaction score SOCNET was still found to have significant effects ($p < 0.01$) in the fricative variation (see Figure 6.8). Participants with more social contact with Jìn speakers were more likely to use the local [x] feature.

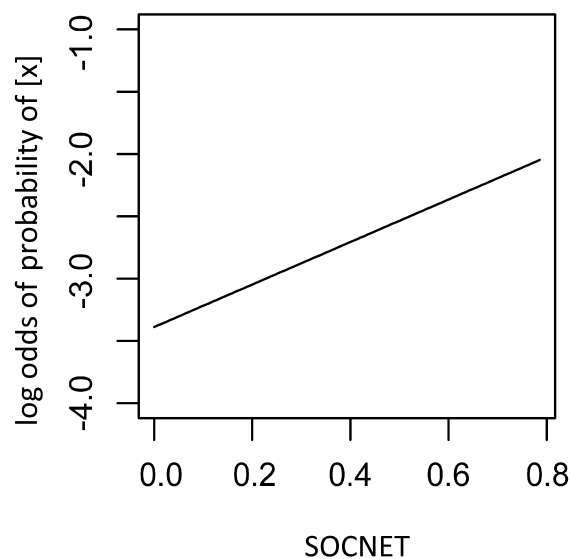


Figure 6.8 Plot of the effect of SOCNET in the New Town data set.

The effects of attitudes were found again in the fricative variable. AGE was interacting with ATTJIN, HOHORE, and ATTOT respectively, indicating that the effects of these three attitudinal index scores are different in the three generations of the New Town speakers, which corresponds to the stress pattern variation results. Figure 6.9 plots the effects of

these interactions. The X-axes are the three attitudinal scores, in which the larger numbers represent more positive attitudes, and the Y-axes are the log odds of the probability of [x] being produced. Again, the expected patterns are positive correlations between attitudes and the production of the fricative variable, which means that speakers with more positive attitudes would show a higher probability of using the [x] variants.

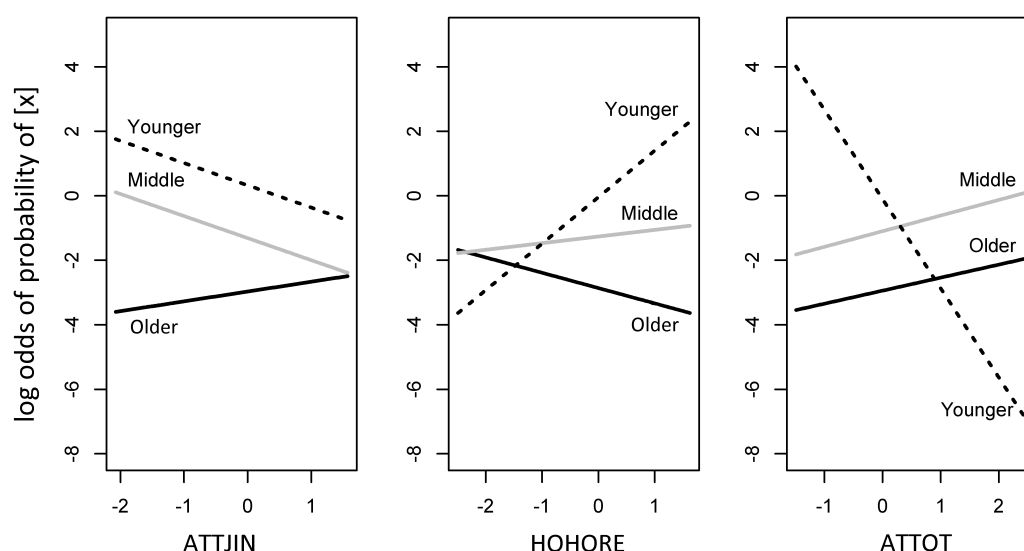


Figure 6.9 Plots of the interaction between AGE and three attitudinal index scores in the New Town data set.

For the older group speakers, there is no significant effect in any of the three attitudinal scores. Only the factor HOHORE showed a near significance trend ($p=0.095$). In the middle group speakers, ATTJIN was found to have a significant effect ($p<0.05$). However, the effect showed that speakers with more positive attitudes towards Jìn dialect were less likely to use the [x] variants, which was contrary to the expected direction. ATTOT was only found to have a near-significance effect ($p=0.068$), and the direction of the effect was an expected positive correlation. For the younger generation, all three attitudinal scores were significantly correlated with their fricative variable production. HOHORE was found to have a very significant effect ($p<0.001$). Speakers who would love to stay in Hohhot were more likely to produce the variants with [x]. As for ATTOT, the same pattern presented in the stress pattern variation result was found here. Speakers with more positive attitudes

towards the Old Town were significantly less likely to use the local [x] variants ($p < 0.001$). A similar negative correlation was also found for ATTJIN (attitudes to Jìn dialect, $p < 0.05$).

To sum up, significant effects of attitudes were also found in the New Town speakers' production of the fricative variable. The general effect patterns were very similar to those found in the stress pattern variable, but there were also some differences. All three attitudinal scores were found to be significant predictors in the younger generation. Some effects still presented negative correlation, but the effect of HOHORE showed the expected direction. ATTJIN was found to have a significant effect in the middle-aged group, but the effect was also a negative correlation. In the older group, the effects of attitudes were not found for the fricative variable. The results are further discussed below.

The younger generation still showed the most interesting results in the fricative variation. Their attitudes towards the Old Town and Jìn dialect were again found to be negatively correlated with the use of Jìn features, which further proved that the effects of ATTOT and ATTJIN found in the stress pattern results were quite robust. The same explanation could be applied here that younger speakers' production of the local Jìn feature might be correlated with their urban Hohhot identity, and speakers with stronger feelings about their Hohhot identity could be more likely to emphasize the distinction between the New Town and the Old Town, thus demonstrating more negative attitudes towards the Old Town and the local dialect. In the stress pattern results, I presented some examples to support the likely changing of the social meaning of the variable, where speakers explicitly related the use of weak-strong pattern with their Hohhot identity. However, since speakers were not explicitly aware of the fricative variation, no one had talked about this feature during the interviews. But the results of the fricative variation analysis here indicated that the fricative variable could be experiencing the same changing process in terms of its social meaning. The younger speakers seemed to produce this feature, perhaps subconsciously, to build their urban Hohhot identity. This could be further supported by the effect of HOHORE found in this variable. Younger speakers' future orientation, which partly reflected their emotional attachment to Hohhot, was also positively correlated with their [x] production. I

said “partly” here because younger speakers’ responses to the statements related to the HOHORE factor were not completely equal to their emotional ties to Hohhot – some speakers who expressed strong feelings about Hohhot in the interview still chose to go to bigger cities in the future for career or education.

In the middle-aged speakers, their attitude to Jìn dialect, ATTJIN, was found to have a significant effect; however, as in the younger generation, the direction of the ATTJIN effect was a negative correlation. This provides evidence for the claim that the new urban Hohhot identity could have already emerged in the second-generation migrants. However, as discussed in section 5.4.1, the ideology is more conservative for the middle-aged speakers because of their struggling identities. They, on the one hand, regarded the label of being a Hohhotian as important, while on the other hand, emphasized their distinction with the Old Town Jìn community. The results of the fricative variation further attested this, as ATTOT, representing their attitudes towards the Old Town also reached a near-significance trend of positive correlation. It is possible that the middle-aged speakers tended to relate the [x] feature with the Old Town communities as well, thus “avoid” using it if they had more negative attitudes towards the local people.

For the older generation, the effects of the three attitudinal index scores were not found in the fricative variation. Only HOHORE was found to have a near-significance effect. However, I do not regard this effect as meaningful. The older group participants were over 70 years old at the time of the interview, so most of them chose to stay in Hohhot not because of their emotional attachment to Hohhot, but because they did not want to move any more at their ages.

To summarize, in the New Town data set, speakers’ attitudinal index scores were again found to have a significant correlation with their fricative variable production. The effects of attitudes are still different in the three generations of migrants. Similarly to the stress pattern variable, the social meaning of the [x] variable could also have changed to be correlated with the newly emerged urban Hohhot identity, and this change might have

already begun in the second generation, but the effect became much stronger in the younger generation.

6.5 Results of the Old Town data set

In the Old Town data set, 940 tokens from 31 speakers were analyzed in binomial mixed effects regression models in R. The linguistic predictors tested were still FVOWEL, PHONEME, and STRESS. The social factors included were AGE, SEX, EDU, MOVNT (whether the speaker has moved to the New Town), the attitudinal index score ATT, and the social interaction score SOCNET. The best-fit model has the fixed effects of FVOWEL, PHONEME, a two-way interaction between AGE and SOCNET, and a three-way interaction between SEX, MOVNT and the attitude score ATT. Speaker was included as a random intercept. The outcome of the best model is shown in Table 6.5, and the plots are shown in Figures 6.10 to 6.12. The results of the linguistic factors were consistent with the New Town results - FVOWEL and PHONEME were both found to be significant predictors of the realization of the [x] variable (see Figure 6.10), whereas STRESS was non significant. Therefore, this section will only focus on the results of the social factors. One thing to note is that the results of the fricative variable here are not totally comparable to the stress pattern variable results, because the stress pattern variation was tested on a subset of non-Jin speakers (mostly younger group speakers), but analysis of the fricative variation included all Old Town participants.

Table 6.5 Output of the best model for the fricative variation in the Old Town data set.

Factors	Estimate	Std.Error	p-value	
(Intercept)	-7.0327	1.7866	8.27E-05	***
FVOWEL _ə	1.8453	0.5264	0.000456	***
FVOWEL _i	-3.2606	0.789	3.59E-05	***
FVOWEL _u	0.5381	0.5328	0.312456	
PHONEME _h	0.9417	0.3193	0.003183	**
PHONEME _p	-0.3995	0.3946	0.311373	
PHONEME _t	0.964	0.355	0.006617	**
SEX _{male}	1.5576	0.4854	0.001332	**
MOVNT _{yes}	1.1017	0.429	0.010234	*
ATT	0.5455	0.709	0.441644	
AGE _{middle}	5.8499	1.8362	0.001443	**
AGE _{younger}	5.2637	1.6562	0.001482	**
SOCNET	6.948	2.1434	0.001189	**
AGE _{middle} * SOCNET	-9.6091	2.4801	0.000107	***
AGE _{younger} * SOCNET	-8.667	2.4297	0.000361	***
SEX _{male} * MOVNT _{yes} * ATT	2.2647	1.0057	0.024335	*

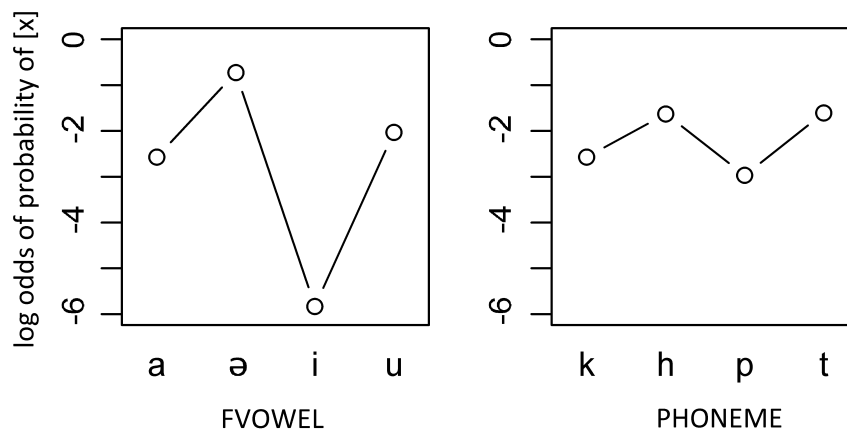


Figure 6.10 Plots of the effects of FVOWEL and PHONEME in the Old Town data set.

The effect of speakers' attitude was found to be significant in the Old Town data set. Similarly to the stress pattern results, the attitudinal score ATT interacted with SEX and MOVNT. The plots of this three-way interaction are shown in Figure 6.11. The X-axes are the attitudinal score ATT representing their general attitudes towards the Old Town, Jin dialect and Hohhot, where a larger number refers to more positive attitudes. The black lines

show the trend of the male speakers, and the gray lines represent the female speakers. The left plot shows the result of speakers who were still living in the Old Town, and the right plot is for those who were born and raised in the Old Town but later moved to the New Town. The female speakers (the gray lines) in both plots showed a trend of positive correlations between their attitudes and production of the [x] variable, but the effects were not statistically significant. The male speakers presented a quite different pattern in terms of their attitudes effects. For males who had moved to New Town (the black line in the right plot), their attitudes had no significant effects on their linguistic production, but for those who were still living in the Old Town (the black line in the left plot), their attitudes showed a significant negative correlation with their fricative variable production ($p < 0.001$).

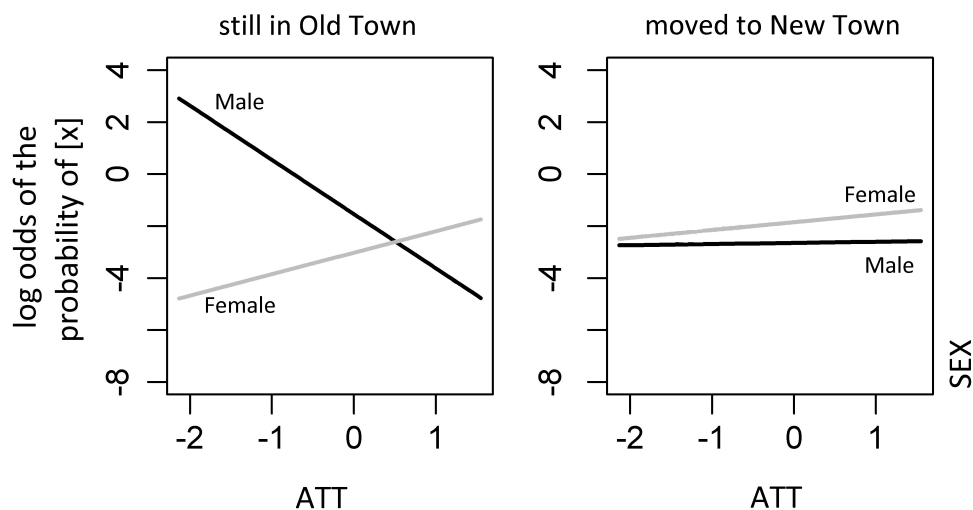


Figure 6.11 Plots of the three-way interaction between ATT, SEX, and MOVNT in the Old Town data set.

These patterns are similar to the stress pattern variation results, in that the female speakers are behaving as expected, but the male speakers are behaving in the opposite way. This could be related to the fact that Old Town male speakers who spoke very local Jìn dialect or showed a lot of Jìn features were mostly middle-aged speakers, who would usually also demonstrate very negative attitudes towards their own accents and the Old Town in the interviews. They often claimed that Jìn dialect was vulgar and unpleasant, but they themselves could not change their accent any more. Some evidence could be found in the

interviews as well. For example, in excerpt (6-2), speaker OMM3, who was a middle-aged male speaker residing in the Old Town, showed very negative attitudes towards Jìn dialect and the Old Town when he was completing the questionnaire, but he was actually speaking a very locally-accented Jìn dialect himself.

(6-2) OMM3, Old Town, middle-aged group, male, Jìn speaker

XW: “Jìn dialect sounds pleasant.” Do you agree?

OMM3: Ah... not pleasant. Disagree! We (Jìn speakers) just can’t change it (Jìn dialect).

OMM3: ...(reading the questionnaire) “I’m proud to be an Old Towner.”...Eh.. how should I put it? Kind of, in the sense of I’m an old Hohhotian, an Old Towner. But now, for the current situation, we are not as good as the other side (New Town).

OMM3: ...(reading the questionnaire) “If my children speak Jìn dialect, I’ll oppose.” Yes, I’ll definitely oppose. I don’t want them to speak Jìn dialect. For their own development, they should speak Pǔtōnghuà to adapt to the society, to the outside world.

On the contrary, male speakers who produced comparatively fewer Jìn features in their speech were mostly younger generation speakers, for whom the distinction between the Old Town and the New Town had become somewhat blurred. Some of them would explicitly express their pride of being an Old Towner, and showed interest in Hohhot traditional culture and the local dialects. This could be the reason a negative correlation was found for the male speakers.

The interaction between AGE and SOCNET was also found to be significant in the best-fit model. The effects of this interaction are presented in Figure 6.12. The SOCNET score is shown on the X-Axis. The three trend lines show the different effects of SOCNET in the three generations of Old Town speakers. The older group speakers showed an expected

pattern, in which people with more Jìn speakers in their social circle showed a higher probability of producing the [x] variants ($p < 0.01$). However, the middle-aged and younger speakers were behaving in the opposite way. Speakers with more Jìn-speaking social networks were less likely to produce the [x] variants (middle-aged: $p < 0.05$; younger: $p = 0.087$, near-significance).

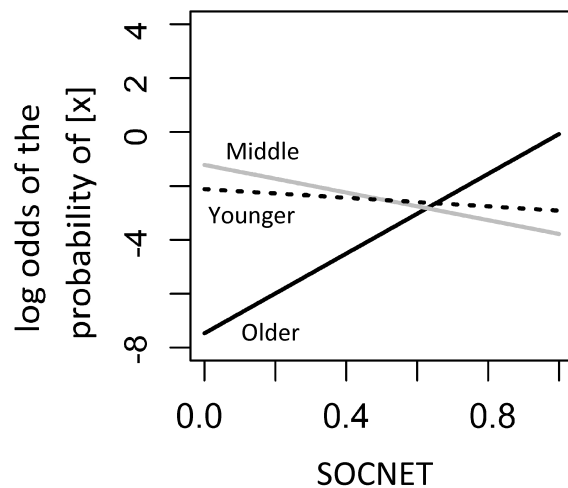


Figure 6.12 Plot of the interaction between AGE and SOCNET in the Old Town data set.

The older speakers were mostly Jìn speakers who were born and raised in the Old Town, so most of them had similar Jìn-speaking family backgrounds. The main issue that contributed to the variation in their social interaction score SOCNET was that some of them had moved to the New Town or received higher education outside Hohhot. So it was reasonable for those who had left the Old Town to converge to the Pǔtōnghuà pronunciation more. For the middle-aged and younger speakers, the situation was different. In the interviews, I found numerous cases where participants with Jìn-speaking family backgrounds were trying to speak in a more standard or formal way, perhaps to show that they were different from their Jìn-speaking families. For example, participant OYF1 had all her family members speaking Jìn dialect, but she herself demonstrated negative attitudes towards Jìn, and was speaking Pǔtōnghuà with few local features involved. In the excerpt (6b), she constantly emphasized that she did not speak any Jìn dialect, and claimed that younger speakers should not speak Jìn.

(6-3) OYF1, Old Town, younger generation, female

OYF1: Yes my contact with Jìn speakers are mainly through my family. And in Hohhot, I think only older people would speak Jìn dialect. The younger generation are usually unwilling to speaker Jìn. And if someone young speaks Jìn, I feel like they've ... like lost their directions (a local expression meaning someone who is aimless or an idler).

Also, middle-aged speakers who had higher SOCNET scores were mostly bi-dialectal speakers, speaking Jìn at home to their family members and speaking Pǔtōnghuà at work places. In the fieldwork, some of the bidialectal speakers tended to make a very clear distinction between Jìn and Pǔtōnghuà, so if they chose to speak Pǔtōnghuà in the interview, they would try to sound very formal and “standard”, which seemed like a sort of hypercorrection. The reason for this could be multifaceted. Some speakers claimed that people should speak either “standard” Jìn dialect or “standard” Pǔtōnghuà, so it was “wrong” or “unpleasant” to use the mixed variety. On the other hand, their hypercorrection could also relate to their identity building in different social contexts. Some speakers might not want to be taken as a Jìn speaker or Old Towner when they chose to speak Pǔtōnghuà. This effect was also found in the stress pattern variation results, in which bi-dialectal speakers with higher SOCNET scores were more likely to be cautious about using Jìn features in their speech. These could be the reasons a negative correlation was found between SOCNET and [x] production in the middle-aged and younger group.

To sum up, in the Old Town data set, speakers' production of the [x] variable was still found to be significantly influenced by their social attitudes. Similarly to the stress pattern results, males and females behaved differently in terms of the effects of their attitudes. And social interaction was also found to be a significant predictor.

From the results of the New Town and Old Town data set, it is plausible to say that speakers' attitudes are very significant predictors of their fricative variable production. Even though the variable is below the level of conscious awareness, speakers may still “unconsciously”

adopt or avoid this feature in different social contexts to build their different identities. The next chapter will compare the results of the stress pattern variable and the fricative variable, and further discuss the role of attitudes in relation to awareness.

7 Discussion

7.1 Attitude as a predictor of linguistic behavior

1. In large-scale dialect contact situations, will speakers' linguistic variation be conditioned by their social attitudes? If so, how will speakers' attitudes influence their linguistic production?
2. Are speakers' explicit attitudes collected by overt questionnaires likely to predict linguistic behavior?

The first two research questions asked in this dissertation directly investigated the connection between speakers' attitudes and their linguistic production. Results from two linguistic features, the stress pattern variable and the fricative variable, both suggested that speakers' use of the local dialect features was significantly predicted by the attitudinal index scores obtained from the AAS questionnaires. In most cases, speakers with more positive attitudes towards Jìn dialect or the Old Town showed a higher probability of producing the Jìn variants. However, the direction of the attitude effects sometimes went the opposite way, which could often be interpreted by the specific social contexts of Hohhot. In the New Town data, the effects of attitudes were found to be different in different age groups, because the three generations of migrants demonstrated different beliefs about the local community and Hohhot, as well as different interpretations about the linguistic features (see more discussion of this in section 7.3). For the Old Town participants, however, the effects of attitudes were different for males and females. Females

with more negative attitudes towards the local dialect tended to linguistically conform to the standard and consciously/unconsciously avoid using Jìn features, but males who negatively evaluated the local dialect were also Jìn speakers themselves or those who retained a lot of Jìn features in their speech. These findings provided evidence for a connection between attitudes and linguistic behavior in the Hohhot contexts.

In this study, I collected speakers' overt or explicit attitudes using self-report questionnaires. However, as discussed in section 2.1.1, previous quantitative studies rarely found a correlation between overt attitudes and linguistic behavior at the individual level. In Gallois, Cretchley, and Watson's (2012) review of direct methods of attitudes measurement like surveys and questionnaires, they pointed out that these measures were excellent indicators of speakers' expressed attitudes, but not always good predictors of behavior (p. 34). The premise of self-report scales of attitude is that people are both willing and able to accurately report their attitudes; however, these conditions are often hard to meet because self-reports of attitudes are highly context dependent (Schwarz, 2008). On the one hand, people may not have introspective access to their attitudes, but need to develop a judgment on the spot, which may depend on what information comes to mind at that point. On the other hand, people may try to hide their attitudes by providing more socially acceptable answers in order to present themselves positively (Bohner & Wanke, 2002).

For the first issue, the AAS questionnaires used in the present study comprised not only direct questions about speakers' attitudes towards Jìn dialect and the Old Town, but also indirect statements that are closely related to these attitudes. Many indirect questions are concerned about their life choices, which prompt the participants to give their judgment on the basis of their experience instead of spontaneous thoughts coming to mind. For example, when I ask the first-generation migrants whether they are willing to be a state-sponsored migrant or whether they are happy to move to the border area, most of them tend to give a positive or neutral answer. However, when the question is presented as "If I were given the choice of staying in Hohhot or going back to my original hometown 30 years ago, I would choose to stay in Hohhot", people tend to report more genuine attitudes because the

question brings them back to a more specific context. Some people even provided their own similar experience about how they made the decision, as they were indeed given the chance of leaving Hohhot when they were young. Similarly, when New Town participants were asked directly about their attitudes towards the Old Town, some of them chose to give a more neutral response and expressed the opinion that one could not judge people according to their origins. However, when they were responding to the statement “If my children/grandchildren are dating someone from Old Town, I would oppose it”, their “true” attitudes were often revealed. Many informants who had encountered similar situations naturally started to talk about their own experience and how they felt when their children were dating an Old Towner (more examples can be found in the excerpts presented in section 5.6). Thus, these indirect questions helped to elicit more genuine attitudes from the speakers. In addition, participants in most interviews were given the chance of elaborating their answers when filling the questionnaire, and I myself, as the interviewer, also encouraged them to talk about these questions based on their own experience. Therefore, participants had sufficient time and opportunities to self-introspect and retrieve more information to form a judgment. As was mentioned in section 4.5.1, there were cases where speakers had second thoughts about their answers after in-depth discussions of related topics and changed their markings on the questionnaire.

The second concern about using consciously offered attitudes is that people may hide their attitudes for social desirability. This issue is often more obvious in the case of face-to-face interviews like the present study compared to self-administered questionnaires (Krysan et al., 1994). For example, Hathett & Schuman (1976, as cited by Schwarz, 2008) found that white informants would mute their negative sentiments about African-Americans when the interviewer was black rather than white. Due to the severe social conflicts between the New Town and Old Town, participants may have avoided showing their negative attitudes if they knew they are talking to someone from the opposite community. However, as an indigenous Hohhotian, my own position and stance in the community helped to elicit more genuine attitudes from the participants. With my paternal and maternal side of the family

being state-sponsored migrants and Jìn speakers respectively, I was able to always stand at the participants' side so that they felt safe to express their true feelings. Moreover, for most participants, I was introduced through their close friends or relatives. At the beginning of each interview, we usually talked about our mutual contacts and our relationship with them so the participants felt closer to me and treated me as a friend as well, which made it easier for me to prompt honest responses from them. Also, due to the age distribution of the participants, most middle-aged speakers and all the older speakers had retired at the time of the interview. Most of them showed great willingness to spend hours chatting with me, especially when they found that I was of a similar age to their children or grandchildren. These factors all contributed to a nice natural context for participants to talk about their true psycho-social orientations.

Apart from the interview contexts, the social situation of Hohhot also made the New Town participants less concerned about expressing their negative attitudes towards the Old Town. The prejudice against the local dialect and locally-born community has been long-standing and prevalent among the migrant community, so it was not hard for them to explicitly convey their negative attitudes. This can also be seen from the online forum posts mentioned in section 3.3.2 and the interview excerpts shown in section 5.6, in which people talked about their prejudice against the Old Town in a very direct manner.

Therefore, the successful use of overt questionnaires in the present study to a large extent increased the validity of the attitudes data collected and thus contributed to the exploration of attitude as a predictor of linguistic behavior.

As discussed in section 2.1.1, another issue with previous studies on the correlation between overtly offered attitudes and linguistic variation is the simplistic measuring and coding process conducted on attitudinal data. In the present study, the use of magnitude continuum in the AAS questionnaires allowed me to gain a fine-grained measurement of the informants' responses to questions related to their attitudes. In dealing with the quantitative data obtained from the questionnaires, the statistical technique PCA was used

to detect multiple aspects of speakers' psycho-social orientations and calculate the relative weight for each question, which better revealed the variation in attitudes between individual speakers. The results further attested to the complex nature of attitudes, indicating that speaker attitude was not as simple as being positive, neutral or negative, but should be treated as a gradient object. For example, although many participants generally demonstrated negative attitudes towards the Old Town, there was still inter-speaker variability in their attitudes, which could potentially contribute to their linguistic variation. Also, PCA revealed four attitudinal factors in the New Town questionnaire responses, but only one factor was revealed in the Old Town data set. This suggests that the different aspects of speakers' attitudes were comparatively more independent for New Town speakers. Their attitudes to Jìn dialect, the Old Town and Hohhot may be different – speakers with very positive attitudes towards Hohhot might at the same time show negative attitudes towards the Old Town and Jìn dialect, which actually corresponds to the complex and struggling identities of the New Town speakers. However, for the Old Town speakers, their attitudes towards Hohhot, the Old Town and Jìn dialect were more unified, with similar directions, because only one attitudinal factor was found.

From the discussion in section 5.6, the attitudinal index scores largely reflect speakers' attitudes presented in the interview, and can be considered as a valid measurement. However, there were also situations where speakers' social orientations were not fully represented by their attitudes, such as in the case of speakers whose attitudes have changed in their lifetime, and those whose attitudes are influenced by other participants in the group settings. Moreover, the questions in AAS could not comprise all aspects of speakers' social attitudes, so there could be more unexplored attitudinal factors having potential effects on speakers' linguistic behavior. For example, as discussed in section 5.4.1, the younger speakers tended to relate the weak-strong pattern to their urban Hohhot identity. These findings further attested to the complex nature of attitudes, especially when exploring long-term language change in dialect contact.

Despite this, the findings of the Hohhot study still suggest that speakers' overt attitudes collected through questionnaires are also likely to predict linguistic behavior. The technique of AAS, as Llamas and Watt (2014) proposed, can be an efficient and useful tool for collecting quick and valid attitudinal data in larger-scale fieldwork, thus providing new angles for future studies of attitude-behavior correlation.

7.2 Interaction between attitudes and social contact

3. Will speakers' attitudes play an independent role in language change if their social contact is also taken into consideration?

4. How will attitudes interact with social contact to influence speakers' linguistic production?

Research question 3 aimed to explore the interaction effects of attitudes and social contact. As discussed in section 2.1.2, the CAT model argues for the effects of both attitudes and interaction in linguistic accommodation, whereas Trudgill's "deterministic model" rejects the effects of attitudes or identity, and advocated for speakers' frequency of contact as the essential factor for speech convergence. However, to test these models in empirical research, it is often difficult to tease apart attitudes from interaction, because their effects are very likely to be highly correlated – speakers with more positive attitudes towards certain people are often more likely to have frequent social contact with them. In the present study, however, both speakers' attitudes and their social interaction with Jìn speakers are finely measured and quantified, which explains more subtle variations between individuals. As a result, the VIF diagnostic method detected no collinearity between the attitudinal index scores and the social networks score in the statistical models, indicating that these factors were not highly correlated, which allowed me to explore the independent effects of attitudes and social contact, as well as the possible interactions between them.

In the results of the two linguistic variables, speakers' social interaction score SOCNET and their attitudinal scores were both found to be significant predictors of their use of the Jìn variants. For the New Town migrant community, people who have more Jìn speakers in their social networks are consistently more likely to use the local dialect features. However, in the Old Town data set, SOCNET scores are sometimes negatively correlated with speakers' use of the two examined Jìn features. This is probably due to the fact that some Old Town participants with Jìn family backgrounds and origins tend to be more concerned about showing Jìn features in their speech to avoid being recognized as Jìn speakers, and some even performed hypercorrection. The findings suggest that speakers' attitudes and their social contact are likely to be independent predictors of linguistic production. The effects of attitudes are still found even when social interaction is considered in the model.

Moreover, when exploring the stress pattern variable, an interaction effect was found between speakers' SOCNET score and their attitudinal index scores in the New Town data. This was especially interesting with regard to research question 4, in which I investigated how attitudes and social contact factors could interact to influence speakers' linguistic behavior. The results showed that for speakers with the same level of social contact, their linguistic production was still predicted by their attitudinal scores, indicating that attitude was playing a role that was independent from the effect of social networks. The finding is consistent with the claims of CAT, and provides a direct counter-argument to Trudgill's interaction-only model. The New Town speakers' adoption of the weak-strong pattern may be not simply a matter of "who interacts with whom" – people with similar frequency of contact with Jìn speakers still displayed linguistic variation that was related to their different psycho-social orientations.

In this interaction effect, it was also found that the effects of attitudes are often stronger or more obvious among speakers with higher SOCNET scores, because speakers who have limited contact with Jìn speakers (or lower SOCNET scores), even if they have very positive attitudes towards the local community, do not have sufficient exposure to the Jìn features and thus will not adopt them to a large extent. Across all speakers, those with both more

positive attitudes and more social contact with Jin speakers often demonstrated the highest probability of using the weak-strong pattern. This indicates that this group of speakers might be leading the change to adopting the weak-strong pattern in the dialect contact situation of Hohhot, probably because they both have the willingness (attitudes) and opportunity (social contacts) to “learn” from the local community.

7.3 The role of attitudes in koinéization

5. What is the role of attitudes in koinéization? Will attitudes play different roles in different generations of speakers in a new dialect formation scenario?

Research question 5 concerns the role of attitudes in different stages of the new dialect formation process. The full data set models exploring the general language change patterns of the two linguistic variables clearly presented the typical koinéization process with a converging and focusing trend between speakers from the New Town and the Old Town. The findings suggested that the focusing happened rapidly in Hohhot in the second generation. The stress pattern variable showed convergence only in the third generation for male speakers, which was because Jin speakers were included in the analysis who used the weak-strong pattern exclusively. Those who varied in their use of the stress pattern, or were bi-dialectal and non-Jin speakers showed convergence with the New Town speakers, presenting a similar probability of using the weak-strong pattern. In the younger generation, the New Town and Old Town speakers showed no significant difference in their use of the two linguistic features, and the convergence became more stable and robust. This rapid focusing process could be explained by the social background of Hohhot to a large extent. The large population of residents with various places of origin and language background led to a highly heterogeneous social situation. The special work-unit social structure formed extremely close-knit social circles in which the contact between migrants and locally-born residents was much more intense and frequent than most of the

koinéization situations found in western societies. As introduced in section 2.2, universal schooling and the formation of peer groups in children are also essential for the speed of focusing (Britain, 1997). In Hohhot, children from the same work-unit often went to the same schools, so most of them were both classmates and neighbors, who naturally formed their own peer groups. Moreover, as the migrants in Hohhot were very much separated from their original hometowns, the second-generation children rarely showed linguistic features from their parents' place of origin, and quickly formed their own way of speaking among their peer groups. These factors could all potentially contribute to the rapid formation of Hū Pǔ.

When exploring the social constraints operating on the two linguistic features in New Town and Old Town speakers separately, speakers' attitudes were found to be a significant predictor of their linguistic production in all models. In the migrant community, the effects of speakers' attitudinal scores were significantly different for the three generations. The reasons behind this could be twofold.

First, it could be explained by the change of social situations in Hohhot for the three age groups. When the first-generation migrants arrived in Hohhot, the New Town and Old Town were more separated, and the linguistic variability was enormous among speakers with different places of origin. The migrants' attitudes towards the local dialect and degree of integration into the local communities were also very different for different individuals. The second-generation migrants grew up at a time when the antagonism between the New Town and Old Town was sharply evident. The linguistic distinction was also obvious between the locally-born Jīn speakers and the migrants who spoke Pǔtōnghuà, so language was taken as an important marker for people's New Town or Old Town identity. However, by the time the younger generation was born, the boundary between the New Town and the Old Town had gradually blurred. Most locally-born people could also speak Pǔtōnghuà, and language was no longer a marker to distinguish the New Town and the Old Town.

Second, the social meaning attached to the examined linguistic variables might have changed for different generations of speakers. The older generation tended to only relate the use of the weak-strong pattern to the local Jìn dialect speakers because they had never heard of this feature before they came to Hohhot. Therefore, they would adopt it to show intimacy with the local speakers or to integrate themselves into the local community. The middle-aged speakers, however, not only related the weak-strong pattern to Jìn speakers, but also seemed to adopt this feature to show their emotional attachment to Hohhot. Therefore, they were found to be struggling with the use of the weak-strong pattern. On the one hand, they might adopt this feature to build their local Hohhot identity, but on the other hand, they understood that this feature was from Jìn dialect and emphasized their distinction from the Old Town speakers, thus they claimed not to use it in their metalinguistic explanations and self-reports. The younger generation, by contrast, seldom related the weak-strong pattern to Old Town speakers, but began to use it as a marker of their urban Hohhot identity, and some of them even expressed this explicitly in the interview. The fricative variable was below speakers' conscious awareness, so they did not explicitly comment on the social meaning of this variable. However, the statistical results of the fricative variable showed a very similar pattern compared to the stress pattern variable, indicating that this feature could undergo a similar process of change in terms of its social meaning. But in any case, to further test the social meaning attached to these variables, perceptual experiments should be conducted in future research.

The findings further attested to the role of psycho-social factors like attitudes and identity in the formation of a new dialect. Attitudes were found to be significant predictors of individuals' linguistic variation in all three generations of speakers, which provide further counter-arguments against Trudgill's deterministic model.

6. Are speakers' attitudes and identities important in shaping the outcome of koinéization?

My answer to research question 6 is somewhat more tentative. From the discussion in section 3.1 and section 4.4.1, Hū Pǔ speakers rarely adopt the Jìn vowels and tones in l-words. When producing an l-word adopted from Jìn dialect, they tend to convert the vowels and tones to the Pǔtōnghuà forms. This is probably due to the salience of the segmental features in Jìn dialect. In Hohhot, if someone produces a word with Jìn dialect vowels or tones in it, he/she is often immediately recognized as a Jìn speaker. This is probably because these features are often seen in Jìn speakers acquiring Pǔtōnghuà as a second dialect; thus they become more salient to non-Jìn speakers. For example, in excerpt (7-1), speaker NMM2, who was a taxi driver, provided an interesting anecdote. When he had two passengers who dressed neatly and looked posh, he expected them to speak Pǔtōnghuà. However, when one of them produced a Jìn vowel in the speech, NMM2 recognized that she was a Jìn speaker at once.

(7-1) NMM2, New Town, middle-aged, male

NMM2: I'm a taxi driver. Once I had two passengers. They were two girls who looked very pretty and posh, and dressed fashionably. Then they began to chat, and later, one of them said, "I really like Hohhot, especially its [tɕ^hũŋ⁵⁵tiɛn⁵⁵] (春天 'spring', pronounced with a Jìn vowel)!" Once I heard this, I was like... oh shit!

The word 春 is pronounced as [tɕ^hun⁵⁵] in Pǔtōnghuà, but the girl in excerpt (7-1) produced it as [tɕ^hũŋ⁵⁵], because [un] merges with [ũŋ] in Jìn dialect. So even though she was speaking Pǔtōnghuà, NMM2 recognized her as a Jìn speaker immediately after she produced this Jìn vowel.

As mentioned in section 2.2, the leveling process in new dialect formation often rules out demographically minority or linguistically marked forms. Kerswill and Williams (2000) in their "principles of koinéization" also claim that, in the outcomes of post-contact varieties, marked regional forms are disfavored (p. 84). This is justified in the Hohhot contexts – vowels and tones that are marked Jìn features are rarely adopted by Hū Pǔ speakers.

However, results from the current study showed that the weak-strong pattern and the [x] variants were kept in the outcome of Hū Pǔ, with possible new social meanings attached to them, which could be considered as a reallocation process (as introduced in section 2.2). The l-words produced by Hū Pǔ speakers are often in an intermediate form, with Pǔtōnghuà vowels and tones, but presenting possible local features in other linguistic levels, like the stress pattern variable and the fricative variable. This intermediate form may correspond to the struggling identities of the New Town migrants. In order to be distinguished from the local Jìn speakers, they never adopted the marked Jìn features like vowels and tones; but to demonstrate their local identity as a Hohhotian, the weak-strong pattern and the [x] variants may be adopted. As a result, the intermediate forms of l-words became a “safer” choice between speakers’ struggling identities; thus they were preferred by Hū Pǔ speakers and became more stabilized in the focusing process. This also implies that the weak-strong pattern and the [x] variants are only possible to be markers of speakers’ Hohhot identity when they are combined with Pǔtōnghuà vowels and tones in the intermediate forms of l-words. If these features are combined with Jìn vowels and tones, they would still be simply taken as pure Jìn features (see more discussion of this in section 7.4).

Therefore, it can be seen from the discussion above that the outcome of l-words forms in Hū Pǔ is very likely to relate to speakers’ social attitudes and their identity construction. However, to examine the full picture of the outcome of Hū Pǔ, more linguistic features in a broader linguistic environment should be further investigated.

7.4 Effect of attitudes in relation to awareness in speech production

7. How will the effects of attitudes be related to the level of awareness of the linguistic variable itself? Is explicit awareness a threshold for attitudes to play a role in linguistic production?

The last research question of this dissertation explores the effects of attitudes in relation to speakers' explicit awareness of the linguistic feature. The two linguistic features selected in this study are subject to difference in their level of awareness. The stress pattern variable is a comparatively salient feature in Jin dialect, whereas the fricative variable is below the level of speakers' conscious awareness. The difference between these two linguistic features in terms of their level of awareness could be addressed from the following perspectives.

The first difference lies in the methodological aspect. At the phase of designing the research, the stress pattern variable was selected as a feature of interest. The elicitation task was designed with a lot of target l-words that could potentially vary in stress pattern. In the interview, when completing the elicitation task, speakers' attention was more on which word they would choose to describe the picture than the attention paid to the actual pronunciation of the word. However, in the later stage of the task, when I began to ask them about their use of l-words, and particularly about the stress pattern choice, speakers' awareness of this feature could be further raised as well as their amount of attention paid to the stress pattern. By contrast, the fricative variable was not considered in the design of the research because even I was not aware of the feature at that time. In the interview and the word elicitation task, this feature was never mentioned by me or by any of the participants, so speakers' production of the fricative variable had little attention paid to it.

Secondly, the stress pattern variable is more likely to show style shifting. As most l-words do not have written forms, I was not able to test speakers' production in a reading style, such as in a word list or passage. However, some participants did report their style shifting

in stress pattern in the metalinguistic comments. For example, NYM5 was a high school teacher in Hohhot. In excerpt (7-2), he reported that he would shift between weak-strong and strong-weak stress patterns according to the formality of the contexts, and he would use strong-weak patterns especially in class. Another speaker OYF2 was a postgraduate student at a university in Beijing, and she was interviewed when she was in Hohhot during her summer vacation. In excerpt (7-3), she claimed that she would use weak-strong and strong-weak patterns interchangeably according to whether she was in Beijing or Hohhot, and with different interlocutors.

(7-2) NYM5, New Town, younger group, male

XW: Do you use /k^hu⁵⁵luŋ⁰/ (s-w) sometimes?

NYM5: hm...in formal contexts I will use /k^hu⁵⁵luŋ⁰/ (s-w), but if it's just daily communication, or when I hang out with friends, I'll say /k^huə⁷⁴³luŋ⁵¹/ (w-s), since everyone can understand it.

XW: What about when you're teaching?

NYM5: Ah, of course in class you can't speak in a too dialectal way. Like the word /k^hu⁵⁵luŋ⁰/ (s-w), if you say "Look! Here's a /k^huə⁷⁴³luŋ⁵¹/ (w-s)" in class, the students will laugh. You should say "Here's a /k^hu⁵⁵luŋ⁰/ (s-w)".

(7-3) OYF2, Old Town, younger group, female

XW: Do you use /k^huə⁷⁴³luŋ⁵¹/ (w-s) or /k^hu⁵⁵luŋ⁰/ (s-w)?

OYF2: I would use /k^huə⁷⁴³luŋ⁵¹/ (w-s).

XW: But what if when you're in Beijing?

OYF2: Of course I'll say /k^hu⁵⁵luŋ⁰/ (s-w) in Beijing. ... But when I met my high school classmates (from Hohhot) in Beijing, I immediately changed back (to w-s).

...

XW: Do you use /xua³⁵la⁰/ (s-w)?

OYF2: No. /xua³⁵la⁰/ (s-w) is Pǔtōnghuà, I'll say that in Beijing.

Therefore, speakers might show style-shifting in the stress pattern choice, and they are likely to have explicit awareness of this according to the self-reports. However, whether the fricative variable will show style-shifting is still unknown since it was difficult to test in l-words, and speakers did not explicitly talk about this feature. However, since the [x] variable is not only found in l-words, future research could test its style-shifting in a wider set of environments.

Thirdly, evidence could be seen from the interview that the stress pattern variable may attract overt social commentary from speakers. As is shown in section 5.4.1, some speakers relate the weak-strong variant to incorrectness, working class and Old Town residents, while the younger generation tends to relate it to local identity, and some even explicitly comment on this. The fricative variable was not mentioned in the interview, so I cannot make a comparison of them in this aspect. However, all participants were given the chance to talk about the features of Jìn dialect in the interview, but none of them mentioned the [x] feature. This also to a certain extent implies that this feature is comparatively less salient. The fact that, I myself, as an indigenous Hohhotian and a linguist, even had no explicit awareness of this feature could be considered as another important piece of evidence for this claim.

Lastly, the stress pattern variable is also more salient because the weak-strong pattern and strong-weak pattern are phonologically very different. An outsider who has no knowledge about Jìn dialect would have difficulty in understanding l-words with weak-strong patterns. By contrast, for the fricative variable, even when [p^h, t^h, k^h, h] are realized as [p^x, t^x, k^x, x], an outsider could still understand them with no effort, because this variable is not involved in maintaining phonological contrasts, thus is less likely to be noticed.

Judging from these clues, the stress pattern variable could be considered as a “stereotype” in Labov’s criterion, or “the third-order indexical” in Silverstein’s order of indexicality (as introduced in section 2.3). In the previous section, I also mentioned the marked Jìn features like vowels and tones in l-words, which are obviously also “stereotypes”, because speakers make direct negative comments on these features. However, although both being stereotypes, the stress pattern variable is not as salient as the marked Jìn features like vowels and tone. Speakers who produced a Jìn vowel may immediately be recognized as a Jìn speaker. But for the stress pattern variable, since the weak-strong pattern has been widely adopted by Hū Pǔ speakers, one cannot make a very confirmative judgment about the linguistic background of the speaker who used it. Therefore, if the comparative salience of these linguistic variables in l-words is considered, as proposed by Honeybone and Watson (2013), the order should be: fricative variable < stress pattern variable < vowels and tones. Since the stress pattern feature and the vowels/tones features are behaving differently in the outcomes of l-words in Hū Pǔ, the findings of this study further support Honeybone and Watson’s claim that salience is a gradient. Also, in the previous section, I also tentatively argued that the possible link between the stress pattern variable and speakers’ urban Hohhot identity is only likely to be found when the weak-strong pattern is in conjunction with Pǔtōnghuà vowels and tones, which attracted speakers’ overt social commentary. But if the weak-strong pattern is combined with Jìn vowels and tones, it is considered a pure Jìn feature, and is less salient than the vowels and tones. This further attests to the complexity of salience as proposed in Campbell-Kibler (2011) and Kristiansen et al. (2011) (see section 2.3). The salience of a linguistic feature is likely to be dynamic and vary in different linguistic and social environments.

With respect to the interplay between attitudes and awareness, this thesis asks whether speakers’ explicit awareness of a linguistic feature is a prerequisite for attitudes to play a role in their adoption of this feature. For the two linguistic features examined in this dissertation, which are greatly different in their level of awareness, speakers’ attitudes were both found to be significant predictors. Even though the speakers were not explicitly

aware of the fricative variable, their production of this feature was still found to be correlated with their psycho-social orientations. The findings imply that explicit awareness is not a threshold for the attitude-language correlation, and thus support Drager & Kirtley's (2016) argument that speakers are likely to unconsciously adopt certain linguistic features as a result of identity construction.

This finding reinforces the importance of social-psychological factors like speakers' attitudes and identities in linguistic convergence and language change. One of the key arguments of Trudgill's deterministic model claims that speech convergence is a quasi-automatic process, which is merely a matter of "who interacts most often with whom" (Trudgill, 2008, p. 251; see section 2.1.2). However, the findings from the present study suggest that even when the linguistic feature is "automatically" or unconsciously adopted, the degree of speakers' adoption of the feature is still likely to be affected by their attitudes.

Drager and Kirtley (2016) explain this sort of unconscious identity construction using exemplar theory. Exemplar theory (Johnson, 1997; Pierrehumbert, 2003; Docherty & Foulkes, 2014) is an experience-based model, arguing that experiences are encoded in the mind as stored episodic memories, or *exemplars*. Phonetic exemplars like the acoustic details of a lexical item, social exemplars like speakers' age and sex, and contextual exemplars related to the conversational event are stored together and form what is called an exemplar cloud. In the fricative variable case, for, example, Hū Pǔ speakers might have experience with hearing local Jìn speakers and very "locally-oriented" migrants producing the [x] variants, so they stored this phonetic information in the exemplar cloud together with the concepts like "localness", "Jìn-related", and "Hohhot". Therefore, when the social representations associated with "localness" were activated, the activation automatically spread to the indexed phonetic exemplars. So the realization of the [x] variants would arise without any need for awareness of the feature. The exemplar theory could be considered as a very good model to explain the findings of the fricative variable in the present study, however, to make firm conclusions, more perceptual experiments should be conducted to

further investigate the application of exemplar theory on the interplay between attitudes and awareness, which will be a very interesting direction for future research.

7.5 The stress pattern variable and the fricative variable

The present study examined the effects of attitudes on two linguistic variables: the stress pattern variable and the fricative variable. The findings suggest that the patterns of the attitudinal effects are generally very similar for the two features, however, they are not exactly the same. For example, the effect of an attitudinal score could be significant in a certain age group in the stress pattern variable, but it may become insignificant in the same age group for the fricative variable. This difference could have partly resulted from the speakers' different levels of awareness of the two linguistic features, as discussed in the previous section. Speakers' explicit awareness of the feature may possibly boost or cover the effects of attitudes to some degree. However, apart from the level of awareness, the two linguistic features are also different in some other aspects, which could also potentially contribute to the different attitudinal patterns found in the results.

For example, the stress pattern variable is a feature that can only be found in l-words (also in some other Jìn lexemes), whereas the fricative variable is a segmental feature that not only appears in l-words, but also in broader linguistic contexts. The use of l-words could be very limited in natural speech because of its specific meaning, so when producing an l-word, speakers are more likely to pay attention to the linguistic features and some people may even show stylistic “control” in certain contexts. By contrast, the fricative variable could be prevalent in natural speech production, so even if speakers had some degree of awareness of this feature, it would still be difficult for them to attend to it all the time.

Moreover, the weak-strong pattern of l-words is a feature that could only be found in Jìn dialect, and Jìn speakers would use it almost exclusively. The [x] variants, however, are a feature that is also likely to be used by Pǔtōnghuà speakers, but it is just more often seen and noticed in Jìn-speaking areas. Jìn speakers also vary in their use of the fricative variable. Therefore, the results of the Old Town speakers are different in the two linguistic variations partly due to whether the monolingual Jìn speakers and bi-dialectal speakers using Jìn dialect in the interview could be included in the analysis. This difference between the two variables may also contribute to the different levels of salience of them, since the link between the [x] variants and Jìn dialect could be weaker than that between the weak-strong pattern and Jìn dialect.

Also, the fricative variable is more likely to be influenced by its linguistic environment than the stress pattern variable. As shown in the results of the fricative variable, it is significantly correlated with the following vowel and the phoneme. The stress pattern, however, is more independent of the surrounding linguistic context, perhaps because it is perceived more as a part of the lexical item. These factors may also influence the patterns of the attitude effects found in the two linguistic variables, thus they could be potentially interesting for future research.

8 Conclusion

This dissertation is an exploration into the role of speakers' attitudes and identities in speech production and language change involved in a large-scale dialect contact situation. Speaker attitude was investigated at the individual level using overt questionnaires, and its correlation with speakers' production of two linguistic variables was explored in mixed effects models. The results suggest that the attitudinal index scores were significant predictors of speakers' use of both variables. And more importantly, the effects of attitudes were still found even when speakers' social interaction with the other community was taken into consideration – for speakers with the same level of social contact, their speech production was still predicted by their attitudes. These findings support the effects of attitudes in individuals' speech production, and imply that attitudes are likely to play a role in linguistic convergence that is independent from the effects of social interaction. This not only provided counter-arguments to Trudgill's "interaction-only" model, but also expanded our understanding of the role of attitudes and social interaction in larger-scale dialect mixing contexts. Moreover, from the methodological perspective, the findings also suggest that speakers' overtly offered attitudes are also likely to predict their linguistic behavior, and the technique of AAS can be a helpful tool for collecting quick and valid attitudinal information in fieldwork for future researchers.

By presenting the case of the formation of the new variety of Hū Pǔ, this study also investigated the role of attitudes in koinéization, or more specifically, the new town scenario of new dialect formation. Due to the intense dialect contact situation between the migrant community and the locally-born community in Hohhot, the new koiné Hū Pǔ

presented a very rapid development trend – the two linguistic variables both showed focusing in the second generation, when New Town and Old Town speakers reached a similar probability of using these features. Investigation of the New Town community found that the attitudinal index scores representing different aspects of their psycho-social orientations had significantly different effects in the three generations of migrants, which could be explained by not only the specific social changes in Hohhot, but also the likely change of social meaning attached to the linguistic variables for speakers of different generations – the adopted local dialect features seemed to be used by later generations as markers of their urban Hohhot identity. Additionally, it was found that when Hū Pǔ absorbed lexemes from the local Jìn dialect, the marked Jìn features like vowels and tones were leveled out, but the two linguistic features examined in this study, the weak-strong pattern and the [x] variants, were kept through reallocation and were refunctionalized as possible markers of the local Hohhot identity. These findings provided further evidence for the role of socio-psychological factors in different stages of new dialect formation and in shaping the outcome of the mixing variety. The situation of Hohhot is not a *tabula rasa* scenario as proposed in Trudgill's deterministic model. However, the findings suggest that the role attitudes and identities play in new dialect formation could be more complex and subtle than we had expected. Therefore, understanding the specific social contexts and speakers' psycho-social orientations should be considered a crucial part in any research of this kind.

The two linguistic features studied also demonstrated different levels of awareness among speakers in Hohhot. Participants overtly commented on the stress pattern variable and reported their style-shifting experience, but they showed no explicit awareness of the fricative variable. However, the effects of attitudes were found on both linguistic features, indicating that speakers' explicit awareness of the feature was not a prerequisite for attitudes-language correlation. Speakers were likely to adopt features unconsciously as a result of identity construction. This finding further supported the effects of attitudes in

speech convergence - even when speakers were adopting features automatically or unconsciously, attitudes were still likely to affect the degree of their adoption.

8.1 Limitations and future research

The linguistic variation examined in this dissertation focused on a set of lexemes: l-words. The findings suggest that l-words contained different levels of linguistic features and the numerous variabilities in these features led to interesting sociolinguistic results. However, due to the limited use of l-words in natural speech, they may not be able to reflect the full picture of the new variety of Hū Pǔ. Therefore, in order to further understand the outcome of the koinéization process in Hohhot, future research should investigate more linguistic variables in a broader linguistic context, like the segmental features and grammatical features absorbed from Jìn dialect.

Also, when exploring these linguistic features, the present study mainly concentrated on inter-speaker variations. However, speakers are also likely to show style shifting in the stress pattern variable and the fricative variable in different contexts. Most l-words do not have written forms, which makes it difficult to compare conversation and reading styles. As discussed in section 7.4, speakers did report their style shifting of the stress pattern variable in the interview. Therefore, it will be interesting to test how speakers shift between weak-strong patterns and strong-weak patterns according to, for example, the formality of contexts, different interlocutors, the presence of different auditors/overhearers, and/or conversation topics. The fricative variable is actually not only present in l-words, so future research can test the reading and conversational style of this variable in a broader linguistic context, which may also justify the level of awareness of this feature. Moreover, the acoustic features of the [x] variants are also potentially interesting for future research.

The attitudinal data in this study were collected in a direct method through the AAS questionnaires. The methodology was proved to be successful in gathering speakers' overt attitudes, but it could not comprise all aspects of their socio-psychological orientations. The use of the PCA technique effectively reduced the dimension of the questionnaires and revealed underlying aspects of speakers' attitudes. However, the PCA process also dropped a great deal of information that could potentially contribute to the variation in the collected attitudes. Moreover, in the model fitting process, more attitudinal information was eliminated due to the problem of multicollinearity, like the IDMIG factor. Therefore, more potentially interesting variations could be contained in the dropped information, which could be further considered in a future direction. Apart from these, speakers' implicit attitudes could also be explored in comparison with the explicit attitudes results, as well as their possible link to speech production.

Since the wedding invitation task was not successfully conducted in the fieldwork, the social networks index score for each participant was calculated on the basis of the self-reports of the proportion of Jìn speakers in their different social circles. The SOCNET scores to a large extent reflected speakers' possible social interaction with Jìn speakers. However, strictly speaking, for two participants who had the same SOCNET score, or the same proportion of Jìn speakers in their social networks, it was still possible that they had different frequency of contact with their Jìn friends or relatives. Moreover, the density of their social network structure was not considered in the SOCNET score either. Therefore, to make stronger arguments about the interaction between attitudes and social interaction, methodologies for collecting more accurate and delicate social networks data could be applied or developed in future research.

When exploring the effects of attitudes in different generations of the migrant community, I argued that the social meaning attached to the adopted linguistic features could have changed. However, this claim was only based on the production results and evidence from the interviews. In order to confirm the indexical meaning of these linguistic features for

different generations of speakers, perceptual experiments should be conducted in future studies to test listeners' reactions to these variables.

One last point is about the participants' distribution. As discussed in section 4.1, due to the limit of my own social networks, the informants in this study may not be perfect representatives of the global population in Hohhot. An example was the Old Town older group participants, many of whom were retired professors from a university in the New Town. Although they were also born and raised in the Old Town, their language and attitudes could both have changed in the New Town environment. Moreover, as my own social networks were highly related to universities, many participants in this study were teachers, professors or residents of university neighborhoods, and fewer working class participants were interviewed. Therefore, in future research, a wider population could be examined with more systematic sampling methods.

Overall, this thesis has contributed to the literature on the role of speakers' attitudes in dialect contact contexts. Notwithstanding the limitations addressed above, the exploration into speakers' overt attitudes and closely related social factors, as well as the innovative techniques in collecting and analyzing attitudinal information, has been fruitful. It has also demonstrated the value of viewing long standing issues in variationist sociolinguistics through the lens of non-western localities (cf Labov, 2015).

References

- Ajzen, I., & Cote, N. G. (2008). Attitudes and the prediction of behavior. In W. D. Crano & R. Prislin (Eds.), *Attitudes and attitude change* (pp. 289–311). New York: Psychology Press.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723.
- Akay, E., & Toraman, Ç. (2015). Students' attitudes towards learning English grammar: A study of scale development. *Journal of Language and Linguistic Studies*, 11(2), 67–82.
- Alwin, D. F. (1973). Making inferences from attitude-behaviour correlations. *Sociometry*, 36(2), 253–278.
- Auer, P., Barden, B., & Grosskopf, B. (1998). Subjective and Objective Parameters Determining “Salience” in Long-term Dialect Accommodation. *Journal of Sociolinguistics*, 2(2), 163–187.
- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Baayen, R. H., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412.
- Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics*, 40(1), 177–189.
- Baker, C. (1992). *Attitudes and language* (Vol. 83). Clevedon, UK: Multilingual Matters.
- Bard, E. G., Robertson, D., & Sorace, A. (1996). Magnitude estimation of linguistic acceptability. *Language*, 72(1), 32–68.
- Bardovi-Harlig, K. (1987). Markedness and Salience in Second-Language Acquisition. *Language Learning*, 37(3), 385–407.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3).

- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *Journal of Memory and Language*.
- Bates, D., Maechler, M., & Bolker, B. (2011). lme4: linear mixed-effects models using S4 classes. R package version 1.1-11. Retrieved March 20, 2016 from <http://cran.r-project.org>.
- Bauer, L. (2008). A question of identity: A response to Trudgill. *Language in Society*, 37(2), 270–273.
- Boersma, P., & Weenink, D. (2017). Praat: doing phonetics by computer [Computer program] (Version 6.0.26). Retrieved from <http://www.praat.org/>.
- Bohner, G., & Dickel, N. (2011). Attitudes and attitude change. *Annual Review of Psychology*, 62, 391–417.
- Bohner, G., & Wänke, M. (2002). *Attitudes and Attitude Change*. Psychology Press.
- Borchigud, W. (1996). Transgressing ethnic and national boundaries: Contemporary “Inner Mongolian” identities in China. In Melissa J. Brown, *Negotiating Ethnicities in China and Taiwan* (pp. 160–182). Berkeley, CA: Institute of East Asian Studies, University of California.
- Borenić, V. K. (2011). Attitudes towards English and FL motivation among croatian University business students: Results of a pilot study. In M. Lehmann, R. Lugossy, & J. Horvath (Eds.), *UPRT 2010: Empirical Studies in English Applied Linguistics* (pp. 135–152). Pécs: Lingua Franca Csoport.
- Bortoni-Ricardo, S. M. (1985). *The urbanization of rural dialect speakers : a sociolinguistic study in Brazil*. Cambridge, New York : Cambridge University Press.
- Britain, D. (1997). Dialect contact, focusing and phonological rule complexity: the Koineisation of Fenland English. *University of Pennsylvania Working Papers in Linguistics*, 4(1).

- Britain, D., & Trudgill, P. (1999). Migration, new-dialect formation and sociolinguistic refunctionalisation: reallocation as an outcome of dialect contact. *Transactions of the Philological Society*, 97(2), 245–256.
- Buchizhuroudecaoyuanlang. (2010, November 5). What's wrong with being an Old Towner. *Tianya*. Retrieved from <http://bbs.tianya.cn/post-259-5378-1.shtml>.
- Campbell-Kibler, K. (2008). I'll be the judge of that: Diversity in social perceptions of (ING). *Language in Society*, 37(5), 637–659.
- Campbell-Kibler, K. (2011). Intersecting variables and perceived sexual orientation in men. *American Speech*, 86(1), 52–68.
- Cedergren, H. J., & Sankoff, D. (1974). Variable rules: Performance as a statistical reflection of competence. *Language*, 50(2), 333–355.
- Chao, Y. R. (1935). 中国方言当中爆发音的种类[The types of plosives in Chinese dialects]. *Bulletin of Institute of Historical and Philosophy*, 5(4), 515–520.
- Chao, Y. R. (1968). *中国话的文法*[A grammar of spoken Chinese]. Berkeley: University of California Press.
- Chen, Y. (2013). 浙南蒲城瓯语方言岛语音研究[A phonetic research on the Pucheng Ou dialect island in South Zhejiang] (unpublished master thesis). Fudan University, Shanghai, China.
- Clark, L., & Trousdale, G. (2013). Using participant observation and social network analysis. In M. Krug & J. Schluter (Eds.), *Research Methods in Language Variation and Change* (pp. 36–52). London: Cambridge University Press.
- Clark, L., & Watson, K. (2016). Phonological leveling, diffusion, and divergence: /t/ lenition in Liverpool and its hinterland. *Language Variation and Change*, 28(1), 31–62.
- Cybulska, K., & Borenić, V. K. (2011). Learning Another Language and Motivation to Continue Learning English: Competition or Synergy? In J. Horvath (Ed.), *UPRT 2011: Empirical Studies in English Applied Linguistics* (pp. 147–159). Pécs: Lingua Franca Csoport.

- Davin, D. (1999). *Internal migration in contemporary China*. Basingstoke: Macmillan.
- Didaojiuchengren. (2014, January 27). Guide for answering outsiders' questions – a must read for Hohhot people. Retrieved from <http://www.sonmg.com/thread-40791-1-1.html>
- Docherty, G. J., & Foulkes, P. (2014). An evaluation of usage-based approaches to the modelling of sociophonetic variability. *Lingua*, 142, 42–56.
- Drager, K. (2009). *A Sociophonetic Ethnography of Selwyn Girls' High* (Unpublished doctoral thesis). University of Canterbury, Christchurch, NZ.
- Drager, K. (2011). Sociophonetic variation and the lemma. *Journal of Phonetics*, 39(4), 694–707.
- Drager, K., & Hay, J. (2012). Exploiting random intercepts: Two case studies in sociophonetics. *Language Variation and Change*, 24(1), 59–78.
- Drager, K., & Kirtley, J. (2016). Awareness, salience, and stereotypes in exemplar-based models of speech production and perception. In A. Babel (Ed.), *Awareness and control in sociolinguistic research* (pp. 1–24). Cambridge: Cambridge University Press.
- Duanmu, S. (2007). *The phonology of standard Chinese*. Oxford: Oxford University Press.
- Dunteman, G. H. (1989). *Principal component analysis. Quantitative applications in the social sciences series* (vol. 69). Newbury Park: CA: Sage Publications.
- Eagly, A. H., & Chaiken, S. (1993). *The Psychology of Attitudes*. Harcourt Brace Jovanovich College Publishers.
- Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. London: Sage publications.
- Fox, S. P. (2007). *The demise of Cockneys?: language change in London's traditional East End*. University of Essex, Colchester, UK.
- Gallois, C. (2013). Language attitudes in context. In H. Giles & B. Watson (Eds.), *The social meanings of language, dialect and accent: international perspectives on speech styles* (pp. 170–179). New York, USA: Peter Lang Publishing.

- Gallois, C., Cretchley, J., & Watson, B. (2012). Methodology in intergroup communication: Eclectic approaches to communicating identity. In H. Giles (Ed.), *The handbook of intergroup communication* (pp. 31–43). New York: Routledge.
- Giles, H., & Coupland, N. (1991). *Language: Contexts and consequences*. Milton Keynes: Open University Press.
- Glasman, L. R., & Albarracín, D. (2006). Forming attitudes that predict future behavior: a meta-analysis of the attitude-behavior relation. *Psychological Bulletin*, 132(5), 778–822.
- Gordon, E., Campbell, L., Hay, J., MacLagan, M., Sudbury, A., & Trudgill, P. (2004). *New Zealand English: its origins and evolution*. Cambridge University Press.
- Gordon, M., Barthmaier, P., & Sands, K. (2002). A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association*, 32, 141–174.
- Guo, J. (2007). 语言态度与方言变异: 溧水县城居民语言态度与语言使用情况的简要调查 [Language attitude and dialect variation: a brief survey of language attitude and language use in Lishui town]. *Social Sciences in Nanjing*, 8, 134–140.
- Haddican, B., Foulkes, P., Hughes, V., & Richards, H. (2013). Interaction of social and linguistic constraints on two vowel changes in northern England. *Language Variation and Change*, 25(3), 371–403.
- Hatchett, S., & Schuman, H. (1975). White respondents and race-of-interviewer effects. *The Public Opinion Quarterly*, 39(4), 523–528.
- Haug-Hilton, N. (2010). *Regional dialect levelling and language standards: changes in the Hønefoss dialect* (unpublished PhD thesis). University of York, York, England.
- Hickey, R. (2000). Salience, stigma and standard. In L. Wright (Ed.), *The development of standard English 1300-1800: Theories, descriptions, conflicts* (pp. 57–72). London: Cambridge University Press.
- Hohhot 3rd National Census Office. (1984). *Hohhot Census Materials from 1949 to 1983*.

- Holmes, J., & Kerswill, P. (2008). Contact is not enough: A response to Trudgill. *Language in Society*, 37(2), 273–277.
- Honeybone, P., & Watson, K. (2013). Salience and the sociolinguistics of Scouse spelling: exploring the phonology of the Contemporary Humorous Localised Dialect Literature of Liverpool. *English World-Wide*, 34(3), 305–340.
- Hou, J. (1999). *现代晋语的研究*[*Studies of modern Jin dialect*]. Beijing: Commercial Press.
- Jankowiak, W. (1993). *Sex, death, and hierarchy in a Chinese city: An anthropological account*. New York, NY: Columbia University Press.
- Jankowiak, W. (2009). Well-Being, cultural pathology, and personal rejuvenation in a Chinese city, 1981–2005. In Gordon Mathews & Carolina Izquierdo (Eds.), *Pursuits of happiness: well-being in anthropological perspective* (pp. 147–166). New York & Oxford: Berghahn books.
- Jankowiak, W. (2013). Urban Mongols: the flourishing of multiple identities in a vibrant ethnic community. *Chinese Sociological Review*, 45(3), 53–73.
- Jinchajifangyan. (2011, March 27). I'm from new town area and I dislike Jin dialect. Retrieved from <http://tieba.baidu.com/f?kz=1036348484>
- Jing, S., & Zhu, Y. (2016). Attitude, maintenance, and use of Beijing dialect and Putonghua among younger generation Beijing-born residents. *Journal of Asian Pacific Communication*, 26(1), 32–55.
- Jingxingjiahe. (2011, July 4). What's wrong with being an Old Towner. Retrieved from <http://tieba.baidu.com/p/3112823281?pid=52365561618&cid=0#52365561618>
- Jin, Y. (2008). 呼和浩特方言中的语气词[Modal particles in Hohhot dialect]. *Modern Science*, 20, 180–181.
- Johnson, D. E. (2009). Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass*, 3(1), 359–383.

- Johnson, K. (1997). Speech Perception without speaker normalization: An exemplar model. In K. Johnson & J. W. Mullenix (Eds.), *Talker Variability in Speech Processing* (pp. 145–165). San Diego: Academic Press.
- Johnson, K. (2002). *Acoustic and auditory phonetics*. Wiley.
- Johnstone, B., Andrus, J., & Danielson, A. E. (2006). Mobility, indexicality, and the enregisterment of “Pittsburghese.” *Journal of English Linguistics*, 34(2), 77–104.
- Karlgren, B. (1915-28). *Études sur la phonologie chinoise [Study of Chinese phonology]*. Leyde et Stockholm.
- Kerswill, P. (1994). *Dialects converging: rural speech in urban Norway*. Clarendon Press.
- Kerswill, P. (2002). Koineization and accommodation. In J. K. Chambers, Peter Trudgill, & Natali Schilling-Estes (Eds.), *The handbook of language variation and change* (pp. 669–702). Oxford: Blackwell.
- Kerswill, P. (2010). Contact and new varieties. In R. H. Chair (Ed.), *The handbook of language contact* (pp. 230–251). Wiley-Blackwell.
- Kerswill, P., & Trudgill, P. (2005). The birth of new dialects. In P. Auer, F. Hinskens, & P. Kerswill (Eds.), *Dialect change: convergence and divergence in European languages* (pp. 196–220). Cambridge: Cambridge University Press.
- Kerswill, P., & Williams, A. (2000). Creating a new town koine: children and language change in Milton Keynes. *Language in Society*, 29(1), 65–115.
- Kerswill, P., & Williams, A. (2002). “Salience” as an explanatory factor in language change: evidence from dialect levelling in urban England. In M. C. Jones & E. Esch (Eds.), *Language Change: the interplay of internal external and extra-linguistic factors* (Vol. 86, pp. 81-110). Berlin, New York: De Gruyter Mouton.
- Khan, A. (2006). *A sociolinguistic study of Birmingham English : language variation and change in a multi-ethnic British community* (Unpublished doctoral thesis). University Of Lancaster, UK.

- Kristiansen, T. (2009). The macro-level social meanings of late-modern Danish accents. *Acta Linguistica Hafniensia*, 41(1), 167–192.
- Kristiansen, T. (2011). Attitudes, ideology and awareness. In R. Wodak, B. Johnstone, & P. Kerswill (Eds.), *The SAGE handbook of sociolinguistics* (pp. 265–278). London: SAGE Publications Ltd.
- Kristiansen, T. (2015). The primary relevance of subconsciously offered attitudes: Focusing the language ideological aspect of sociolinguistic change. In A. Prikhodkine & D. R. Preston (Eds.), *IMPACT: studies in language and society* (Vol. 39, pp. 85–116). Amsterdam: John Benjamins Publishing Company.
- Kristiansen, T., Maegaard, J. M., & Pharo, N. (2011). Exploring the social meaning of Copenhagen Danish /s/. Presented at the 6th International Conference on Language Variation in Europe, Friburg.
- Krysan, M., Schuman, H., Scott, L. J., & Beatty, P. (1994). Response rates and response content in mail versus face-to-face surveys. *The Public Opinion Quarterly*, 58(3), 381–399.
- Labov, W. (1963). The social motivation of a sound change. *Word*, 19(3), 273–309.
- Labov, W. (1972). *Sociolinguistic patterns*. Oxford: Blackwell.
- Labov, W. (2001). *Principles of linguistic change, social factors*. Wiley.
- Labov, W. (2015). The discovery of the unexpected. *Asia-Pacific Language Variation*, 1(1), 7–22.
- Ladegaard, H. J. (2000). Language attitudes and sociolinguistic behaviour: Exploring attitude-behaviour relations in language. *Journal of Sociolinguistics*, 4(2), 214–233.
- Ladegaard, H. J., & Sachdev, I. (2006). “I like the Americans... but I certainly don’t aim for an American accent’: language attitudes, vitality and foreign language learning in Denmark. *Journal of Multilingual and Multicultural Development*, 27(2), 91–108.
- Language in society* 37(2). (2008). Cambridge: Cambridge University Press.
- Lan, J. (2012). 呼和浩特话中的人称代词“俄”与“/nia/”[Personal pronouns “é” and “niá” in Hohhot dialect]. *Modern Chinese*, 4, 35–38.

- Lan, Y., Cai, B., & Zuo, F. (2010). 攀枝花市方言研究综述兼谈攀枝花市方言分区[The summarization of the research on Panzhihua dialect concurrently on classification/distribution]. *Journal of Panzhihua University*, 27(2), 34–38.
- Laowannao09. (2011, January 31). What's wrong with being an Old Towner. Retrieved from <http://bbs.tianya.cn/post-259-5378-1.shtml>.
- Lass, R. (1984). *Phonology: an introduction to basic concepts*. Cambridge: Cambridge University Press.
- Lefcheck, J. (2012, December 28). Dealing with multicollinearity using VIFs. Retrieved from <http://jonlecheck.net/2012/12/28/dealing-with-multicollinearity-using-variance-inflation-factors/>
- Le Page, R. B., & Tabouret-Keller, A. (1985). *Acts of identity: creole-based approaches to language and ethnicity*. CUP Archive.
- Liang, D. (1985). 四川省渡口市方言的现状和未来[Current status and outlook of dialects in Dukou, Sichuan]. *Dialects*, (4), 291–296.
- Liu, W., & Liu, X. (2000). *呼和浩特方言辨正 [Hohhot dialect]*. Hohhot: Inner Mongolia People's Press.
- Liu, Y. (2009). 阶层、语体和语言变异: 以昆区北方话鼻韵尾变异为例[Social-status, style and language variation: the case of nasal variation in Kundulun Baotou]. *Social Sciences in Nanjing*, (6), 114–121.
- Liu, Y., & Xu, D. (2012). 移民城市语言变异的本土化: 以包头昆都仑鼻韵尾变异为例[The localization of language variation: with the example of nasal variation in Kundulun, Baotou]. *Studies in Language and Linguistics*, 32(1), 92–100.
- Li, W. (1994). *Three generations, two languages, one family: Language choice and language shift in a Chinese community in Britain* (Vol. 104). Clevedon, UK: Multilingual Matters Ltd.
- Li, Z., & Xin, S. (1987). 呼和浩特方言的一些句法特点[Some syntactic features of Hohhot dialect]. *Journal of Inner Mongolia University (Philosophy and Social Sciences)*, 2, 101–109.

- Llamas, C. (2007). "A place between places": language and identities in a border town. *Language in Society*, 36(4), 579–604.
- Llamas, C., & Watt, D. (2014). Scottish, English, British?: innovations in attitude measurement. *Language and Linguistics Compass*, 8(11), 610–617.
- Lu, F. (1989). The work unit: A special form of social organisation. *Chinese Social Science*, (1), 71–88.
- Lu, X. (2014). 武汉市青山言语社区“弯管子话”研究[A study of the Wanguanzi accent in Qingshan speech community of Wuhan] (unpublished master thesis). Fudan University, Shanghai, China.
- Mæhlum, B. (1992). Dialect socialization in Longyearbyen, Svalbard (Spitsbergen): A fruitful chaos. In E. H. Jahr (Ed.), *Language Contact: Theoretical and Empirical Studies* (pp. 117–130). Berlin: Mouton de Gruyter.
- Ma, G., & Xing, X. (1997). 内蒙古晋语析音词研究[Study of l-words in Inner Mongolia Jin dialects]. In X. Xing & Y. Zhang (Eds.), *Study of western Inner Mongolia dialect grammar* (pp. 251–262). Hohhot, China: Inner Mongolia People Publishing House.
- McKenzie, R. M. (2006). *A quantitative study of the attitudes of Japanese learners towards varieties of English speech: Aspects of the sociolinguistics of English in Japan* (Unpublished doctoral thesis). University of Edinburgh, UK.
- McKenzie, R. M. (2010). *The social psychology of English as a global language: attitudes, awareness and identity in the Japanese Context*. Dordrecht: Springer Science & Business Media.
- Mesthrie, R. (1992). *Language in indenture: a sociolinguistic history of Bhojpuri-Hindi in South Africa*. Routledge.
- Milroy, L. (1980). *Language and social networks*. Blackwell Oxford.
- Mostyn, B. J. (1978). *The attitude behavior relationship*. Bradford: MCB Publications.
- National Bureau of Statistics, NBS (2013). *2010 China census materials*. Retrieved January 5, 2015 from <http://www.stats.gov.cn/tjsj/pcsj/rkpc/6rp/indexch.htm>.

- Nycz, J. (2016). Awareness, salience, and stereotypes in exemplar-based models of speech production and perception. In A. Babel (Ed.), *Awareness and control in sociolinguistic research* (pp. 62–79). Cambridge: Cambridge University Press.
- Omdal, H. (1977). Høyangermålet-en ny dialekt. *Språklig Samling*, (1), 7–9.
- Oppenheim, A. N. B. (1992). *Questionnaire design, interviewing, and attitude measurement* (New ed.). New York: Pinter Publishers.
- O'Reilly, K. (2005). *Ethnographic methods*. Psychology Press.
- Pardo, J. S., Gibbons, R., Suppes, A., & Krauss, R. M. (2012). Phonetic convergence in college roommates. *Journal of Phonetics*, 40(1), 190–197.
- Pierrehumbert, J. B. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech*, 46(2-3), 115–154.
- Podesva, R. J. (2006). Intonational variation and social meaning: categorical and phonetic aspects. *U. Penn Working Papers in Linguistics*, 12(2), 189–202.
- Puthuval, S., & Wang, X. (2017). The language of Hohhot. In R. Sybesma, W. Behr, Z. Handel, & J. Huang (Eds.), *Encyclopedia of Chinese language and linguistics*. (pp. 379–386). Leiden: Brill Academic Publishers.
- Qiao, Q. (2005). 晋方言研究综述[Review of studies on Jìn dialect]. *Journal of Shanxi University, Philosophy and Social Science*, 28(1), 84–89.
- Rácz, P. (2013). *Salience in sociolinguistics: a quantitative approach* (Vol. 84). Berlin: De Gruyter Mouton.
- Rand, D., & Sankoff, D. (1990). *Goldvarb Version 2: A variable rule application for the Macintosh*. Université de Montréal.
- R Core Team. (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, R version 0.98.1103. Retrieved April 26, 2015 from <http://www.R-project.org/>.

- Redinger, D. (2010). *Language attitudes and code-switching behaviour in a multilingual educational context: the case of Luxembourg* (Unpublished doctoral thesis). University of York, York, UK.
- Revelle, W. (2016). *How to: use the psych package for factor analysis and data reduction*. Springer.
- Revelle, W. (2015). Psych: procedures for personality and psychological research. Northwestern University, Evanston, Illinois, USA. R package version: 1.5.8. Retrieved June 15, 2015 from <http://CRAN.R-project.org/package=psych>.
- Rong, X., & Shi, T. (2001). Inequality in Chinese Education. *Journal of Contemporary China*, 26(10), 107–124.
- Røyneland, U. (2005). *Dialektnivellering, ungdom og identitet: Ein komparativ studie av språkleg variasjon og endring i to tilgrensande dialektområde, Røros og Tynset*. (Unpublished doctoral thesis). University of Oslo, Oslo.
- Sandve, B. H. (1976). *Om talemålet i industristaden Odda og Tyssedal: Generasjonsskilnad og tilnærming mellom dei to målføra* (Unpublished doctoral thesis). University of Bergen, Bergen.
- Schilling, M. S. (2013). *Language attitudes of University of Cape Town linguistics students towards codeswitching* (Unpublished doctoral thesis). The College of William & Mary, Williamsburg, US.
- Schneider, E. W. (2008). Accommodation versus identity? A response to Trudgill. *Language in Society*, 37(2), 262–267.
- Schwarz, N. (2008). Attitude measurement. In W. D. Crano & R. Prislin (Eds.), *Attitudes and Attitude Change* (pp. 41–60). New York: Psychology Press.
- Shadle, C. H. (1991). The effect of geometry on source mechanisms of fricative consonants. *Journal of Phonetics*, 19, 409–424.
- Shen, M. (2017). Jìn dialect. In R. Sybesma, W. Behr, Z. Handel, & J. Huang (Eds.), *Encyclopedia of Chinese language and linguistics*. Leiden: Brill Academic Publishers.

- Shepard, C. A., Giles, H., & Le Poire, B. A. (2001). Communication accommodation theory. In W. P. Robinson & H. Giles (Eds.), *The new handbook of language and social psychology* (pp. 33–56). Chichester, UK: John Wiley & Sons, Ltd.
- Siegel, J. (1985). Koines and koineization. *Language in Society*, 14(3), 357–378.
- Siegel, J. (1987). *Language contact in a plantation environment*. Cambridge: Cambridge University Press.
- Siegel, J. (1997). Mixing, leveling, and pidgin/creole development. In A. Spears & D. Winford (Eds.), *The structure and status of pidgins and creoles* (pp. 111–149). Amsterdam: John Benjamins Publishing Company.
- Siegel, J. (2010a). Pidgins and creoles. In N. H. Hornberger & S. L. McKay (Eds.), *Sociolinguistics and language education* (pp. 232–264). Bristol: Multilingual Matters.
- Siegel, J. (2010b). *Second dialect acquisition*. Cambridge University Press.
- Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language & Communication*, 23(3–4), 193–229.
- Solheim, R. (2006). *Språket i smeltegryta: Sosiolingvistiske utviklingslinjer i industrisamfunnet Høyanger* [Language in a melting pot: sociolinguistic lines of development in the Høyanger industrial community] (Unpublished doctoral thesis). Norwegian University of Science and Technology.
- Solheim, R. (2009). Dialect development in a melting pot: The formation of a new culture and a new dialect in the industrial town of Høyanger. *Nordic Journal of Linguistics*, 32(2), 191–206.
- Stausland Johnsen, S. (2015). Dialect change in South-East Norway and the role of attitude in diffusion. *Journal of Sociolinguistics*, 19(5), 612–642.
- Stuart-Smith, J., Pryce, G., Timmins, C., & Gunter, B. (2013). Television can also be a factor in language change: Evidence from an urban dialect. *Language*, 89(3), 501–536.

- Sudbury, A. (2000). *Dialect contact and koineization in the Falkland Islands: Development of a Southern Hemisphere Variety?* (Unpublished doctoral thesis). University of Essex, Colchester.
- Sudbury, A. (2001). Is Falkland Island English a southern hemisphere variety? *English World-Wide*, 22, 55–80.
- Sun, D. (2009). 江汉油田话“潜”字声调变异调查研究[An analysis of the tone variation of Qian(潜) in the dialect of Jiangnan oilfield]. *Studies in Language and Linguistics*, 29(1), 48–53.
- Sun, D. (2011). 语言认同与语言变化: 江汉油田语言调查[Language identity and language change: a survey of language use in the Jiangnan oilfield]. *Applied Linguistics*, (1), 27–37.
- Sun, D. (2013). 工业化过程中的语言变异与变化: 江汉油田言语社区调查[Language change and variation in industrialization: the Jiangnan oilfield study]. China Social Sciences Press.
- Sun, J. (2006). Fission reduplication in modern Chinese dialects. *Bulletin of Chinese Linguistics*, 1(1), 109–131.
- Tagliamonte, S. (2011). *Variationist sociolinguistics: Change, observation, interpretation*. Oxford: Wiley-Blackwell.
- Tang, Z. (2016). Language choice and language attitude among different social segments in Shenzhen. *Journal of Asian Pacific Communication*, 26(1), 143–160.
- Tang, Z., & Li, S. (2016). 湖南城步长安营西南官话方言岛研究[A study of the Chengbu Changanying Southwest Mandarin dialect island]. *Cultural Journal*, 9, 164–167.
- Trudgill, P. (1986). *Dialects in contact*. Oxford: Blackwell.
- Trudgill, P. (2004). *New-dialect formation: The inevitability of colonial Englishes*. Oxford University Press, USA.
- Trudgill, P. (2008). Colonial dialect contact in the history of European languages: On the irrelevance of identity to new-dialect formation. *Language in Society*, 37(2), 241–254.
- Van den Berg, M. (2016). *Special Issue, Journal of Asian Pacific Communication*, 26(1).

- Van den Berg, M., & Xu, D. (2010). *Industrialization and the restructuring of speech communities in China and Europe*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Walder, A. G. (1986). *Communist neo-traditionalism: Work and authority in Chinese industry*. Berkeley: University of California Press.
- Wang, H. (1994). 汉语常用的两种语音构词法: 从平定儿化和太原嵌 l 词谈起[Two common phonetic word-formation in Chinese: Pingding erhua and “l-words” in Taiyuan dialect]. *Studies in Language and Linguistics*, 26, 65–78.
- Wang T., & Ding H. (2012). Production and perception of nasal coda mergence in Hohhot dialect: from a prosodic point of view. *Proceedings of the 3rd international symposium on tonal aspects of languages* (pp. 1-6), Nanjing, China.
- Watson, K., & Clark, L. (2015). Exploring listeners’ real-time reactions to regional accents. *Language Awareness*, 24(1), 38–59.
- Watt, D. (1998). *Variation and change in the vowel system of Tyneside English* (Unpublished doctoral thesis). University of Newcastle upon Tyne.
- Wright, R., Hargus, S., & Miller, J. C. (2008). Using acoustics to resolve place controversies in deg xinag fricatives. *Canadian Acoustics*, 36(3), 66–67.
- Wu, C. (2006). 包头鼻韵尾变异: 言语社区的定量分析[Nasal variation in Baotou: a quantitative analysis of the speech community] (unpublished master thesis). Nanjing University, Nanjing, China.
- Xiaoduantuixiaoyanjing. (2014, June 18). Every time I come to the Hohhot forum, I gain some knowledge. Retrieved from <http://tieba.baidu.com/p/3112823281?pid=52365561618&cid=0#52365561618>
- Xing, X. (1960). 内蒙古晋语的“甚”“咋”及其他[“Shèng” and “zǎ” in Inner Mongolian Jìn dialect]. *Journal of Language and Literature Studies*, 30, 36–37.
- Xing, X. (1995). 内蒙古晋语的语法特点[Grammatical features of Inner Mongolia Jìn dialects]. *Journal of Inner Mongolia University (Philosophy and Social Sciences)*, (1), 52–58.

- Xu, D. (1992). *A sociolinguistics study of Mandarin nasal variation* (Unpublished doctoral thesis). University of Ottawa, Ottawa.
- Xu, D. (2004). 言语社区理论[The theory of speech community]. *The Journal of Chinese Sociolinguistics*, 2(1), 18–28.
- Xu, D. (2006). Urban language survey. *The Journal of Chinese Sociolinguistics*, (2), 1–15.
- Xu, D. (2008). 语言的变异性与言语社区的一致性[Linguistic variability and the uniformity of speech community: A quantitative analysis of nasal variation in Mandarin Chinese]. *Language Teaching and Linguistic Studies*, (5), 78–86.
- Xu, D. (2010). The formation of a speech community: Mandarin nasal finals in Baotou. In M. Van den Berg & D. Xu (Eds.), *Industrialization and the Restructuring of Speech Communities in China and Europe* (pp. 120–140). Newcastle upon Tyne: Cambridge Scholars Publishing.
- Xu, T. (1981). 山西平定方言的“儿化”和晋中的所谓嵌“l”词[Erhua in Pingding dialect and the so-called “l-words” in Jinzhong dialect]. *The Chinese Language*, 6, 408–415.
- Xu, X. (2008). 包头新城区柯因内语的形成与发展[Koineization in new town area of Baotou] (unpublished master thesis). Nanjing University, Nanjing, China.
- Yaeger-Dror, M. (1993). Linguistic analysis of dialect “correction” and its interaction with cognitive salience. *Language Variation and Change*, 5(2), 189–224.
- Yang, J. (1997). 洛阳市普通话和方言的分布与使用情况[The distribution and use of Putonghua and dialects in Luoyang]. *Applied Linguistics*, 24(4), 57–61.
- Yang, J. (1999). 试论中国新型工业区语言状态研究[On the study of the language state in a developing industrial area of China]. *Applied Linguistics*, 29(1), 8–10.
- Yang, J. (2002). 中国新型工业区语言状态研究[Research on language status in new industrial cities in China]. *Linguistic Research*, 82(1 and 2), 13–21 and 28–32.
- Yang, J. (2004). 中国城市语言研究的若干思考[Reflections on the research of China’s urban linguistic varieties]. *The Journal of Chinese Sociolinguistics*, 2(1), 46–53.

- Yang, L. (2010). 言语社区理论和武汉市青山区武钢的语言调查[Speech community theory and language survey in Qingshang District of Wuhan]. *Huazhong Humanity Forum*, (2), 78–81.
- Yang, W. (2013). 江西上饶铁路话调查研究[Research on Shangrao railway dialect in Jiangxi] (Unpublished doctoral thesis). Fudan University, Shanghai, China.
- Yanyanaichitang. (2013, May 9). On hearing this, I laughed. Retrieved from <http://tieba.baidu.com/p/2315821480?pid=32500087877&cid=0#32500087877>
- Yayayaya. (2010, November 29). What's wrong with being an Old Towner. Retrieved from <http://bbs.tianya.cn/post-259-5378-1.shtml>.
- Yu, A. C. L., Abrego-Collier, C., & Sonderegger, M. (2013). Phonetic imitation from an individual-difference perspective: Subjective attitude, personality and “autistic” traits. *PloS ONE*, 8, 1–13.
- Yuchenshang. (2014, June 18). What's wrong with Nanchafang? Are you all really that good? Retrieved from <http://tieba.baidu.com/p/3112100630?pn=2>
- Yun, M. (2017). Rù 入 tone development in non-Mandarin dialects. In R. Sybesma, W. Behr, Z. Handel, & J. Huang, *Encyclopedia of Chinese language and linguistics*. Leiden: Brill.
- Zhao, B. (1979). 晋中话“嵌 l 词”汇释[“L-words” in Jinzhong dialect]. *The Chinese Language*, 6, 455–458.
- Zhou, M. (2010). Sociolinguistic research in China. In M. J. Ball (Ed.), *The Routledge handbook of sociolinguistics around the world* (pp. 67–80). London and New York: Routledge.
- Zhou, W. (2011). 语言态度和语言使用的相关性分析: 以 2007 年南京城市语言调查为例 [The relationship between language attitudes and language uses: Taking 2007's Nanjing urban language survey for example]. *Language Survey and Research*, (1), 89–96.
- Zhu, X. (2010). 北方话鼻韵尾变异: 包头昆都仑区 16 年后的再调查[Nasal rime variation of northern dialects: a reinvestigation of Kundulun District of Baotou after 16 years]. *Private Education Research*, 55(4), 62–68.

Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, 1(1), 3–14.

Appendix A: List of L-words

The 16 target l-words studied in this thesis are presented below.

	L-words	Jìn dialect	Pǔtōnghuà
1	骨碌 ‘roll’	/kuəʔ ⁴³ lu ³⁵ /	/ku ⁵⁵ lu ⁰ /
2	趿拉 ‘wear shoes like slippers’	/t ^h əʔ ⁴³ la ³¹ / or /t ^h əʔ ⁴³ la ⁰ /	/t ^h a ⁵⁵ la ⁰ /
3	扑拉 ‘push slightly to remove the dirt or make soft things more smooth’	/p ^h əʔ ⁴³ la ³¹ /	/p ^h u ⁵⁵ la ⁰ /
4	划拉 ‘scribble’	/xuəʔ ⁴³ la ³⁵ /	/xua ³⁵ la ⁰ /
5	提溜 ‘carry sth. randomly’	/tiəʔ ⁴³ liu ³¹ / or /t ^h iəʔ ⁴³ liu ³¹ /	/ti ⁵⁵ liu ⁰ /
6	卜烂 ‘trip up, stumble’	/pəʔ ⁴³ lǎ ³⁵ / or /p ^h əʔ ⁴³ lǎ ³⁵ /	NA
7	拨拉 ‘repeatedly move horizontally with hand or stick’	/pəʔ ⁴³ la ³¹ / or /p ^h əʔ ⁴³ la ³¹ /	/pu ⁵⁵ la ⁰ /
8	圪料 ‘bended, curved’	/kəʔ ⁴³ liəu ³⁵ /	NA
9	耷拉 ‘droop, hanging’	/təʔ ⁴³ la ³¹ / or /t ^h əʔ ⁴³ la ³¹ /	/ta ⁵⁵ la ⁰ /
10	嘟噜 ‘a cluster/bunch of sth.’	/tuəʔ ⁴³ lu ³¹ / or /t ^h uəʔ ⁴³ lu ³¹ /	/tu ⁵⁵ lu ⁰ /
11	合浪 ‘an alley or a lane’	/xəʔ ⁴³ lǎ ³⁵ /	NA
12	葫芦 ‘a calabash’	/xuəʔ ⁴³ lu ³⁵ /	/xu ³⁵ lu ⁰ /
13	轱辘 ‘wheels’	/kuəʔ ⁴³ lu ³⁵ /	/ku ³⁵ lu ⁰ /
14	卜浪 ‘an onomatopoeia for the sound of a rattle-drum’	/pəʔ ⁴³ lǎ ³⁵ /	/po ⁵⁵ laŋ ⁰ /
15	陀螺 ‘a spinning top’	/t ^h uəʔ ⁴³ luo ³¹ /	/t ^h uo ³⁵ luo ³⁵ /
16	窟窿 ‘holes’	/k ^h uəʔ ⁴³ luŋ ³⁵ /	/k ^h u ⁵⁵ luŋ ⁰ /

The following l-words are not target words of the elicitation task, but are produced by participants during the interview.

	L-words	Jìn dialect	Pǔtōnghuà
17	胡落 ‘push slightly to make soft things more smooth’	/xuəŋ ⁴³ lu ³⁵ /	/xu ³⁵ luo ⁰ /
18	当啷 ‘hang, slouch’	/təŋ ⁴³ lǎ ³⁵ /	/taŋ ⁵⁵ laŋ ⁰ /
19	出溜 ‘slide, slip’	/tɕ ^h uəŋ ⁴³ liu ³¹ /	/tɕ ^h u ⁵⁵ liu ⁰ /
20	卜浪 ‘a stick’	/pəŋ ⁴³ lǎ ³⁵ /	NA
21	出链 ‘a string of something’	/tɕ ^h uəŋ ⁴³ liɛ ³⁵ /	NA
22	得老 ‘head’	/təŋ ⁴³ lao ³¹ /	NA
23	仄愣 ‘standing, sticking out’	/tsəŋ ⁴³ ləŋ ³⁵ /	/tɕŋ ⁵⁵ ləŋ ⁰ /
24	圪揽 ‘a pole or a rod’	/kəŋ ⁴³ lǎ ³¹ /	NA
25	圪佬 ‘take sth. out from under the couch/bed’	/kəŋ ⁴³ lao ³¹ /	NA
26	圪链 ‘a squirrel’	/kəŋ ⁴³ liɛ ³⁵ /	NA
27	圪溜 ‘bended, curved’	/kəŋ ⁴³ liu ³⁵ /	NA
28	骨碌 ‘a section of sth. like a round stick’	/kuəŋ ⁴³ lu ³⁵ /	/ku ⁵⁵ lu ⁰ /
29	合喇 ‘a gap or crack, often on the furniture’	/xəŋ ⁴³ la ³⁵ /	NA
30	忽陆 ‘get muddled or silly’	/xuəŋ ⁴³ lu ³⁵ /	/xu ³⁵ t ^h u ⁰ /
31	囫囵 ‘whole, entire’	/xuəŋ ⁴³ luɛ ³⁵ /	/xu ³⁵ lun ³⁵ /
32	机灵 ‘clever, smart, often referring to kids’	/tɕiəŋ ⁴³ liu ³⁵ /	/tɕi ⁵⁵ liŋ ⁰ /
33	摸捞 ‘touch’	/məŋ ⁴³ lau ³⁵ /	NA
34	筐箩 ‘a wicker basket’	/p ^h əŋ ⁴³ luo ³¹ /	/p ^h o ²¹ luo ⁰ /
35	扑捞 ‘push repeatedly to remove dirt’	/p ^h əŋ ⁴³ lau ³¹ /	NA
36	突隆 ‘pants legs hang loose, to describe someone messy’	/t ^h uəŋ ⁴³ luŋ ³¹ /	/t ^h u ⁵⁵ lu ⁰ /

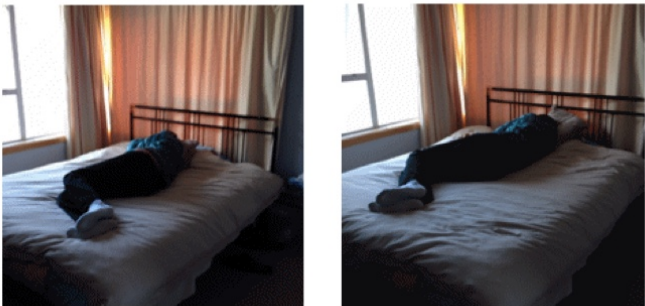
Appendix B: Slides used in the word elicitation task

(Note: Some of the pictures are animations. The slides can be downloaded at:

<https://drive.google.com/file/d/0B8TouWJ94EU4dEZKQkdUQ2ZlVzg/view?usp=sharing>)

Slide 1 Target l-word: 骨碌 ‘roll’


下面两个动图有什么区别？
What is the difference between the two pictures?



The slide contains two side-by-side photographs of a person lying on a bed. The person is wearing dark clothing and is positioned on their side. The bed has a white sheet and a dark headboard. The background shows a window with curtains. The two photographs are identical, showing the person in the same position on the bed.

Slide 2 Target l-word: 趿拉 ‘wear shoes like slippers’

下面三个动图有什么区别？
What is the difference between the three pictures?

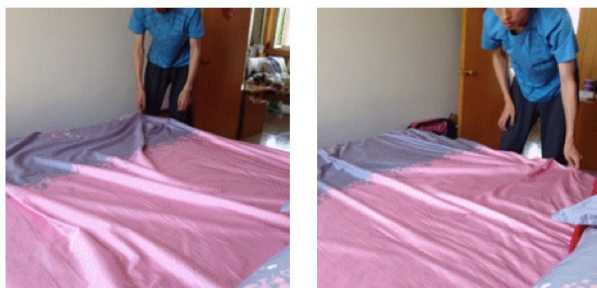


The slide contains three side-by-side photographs showing a person's feet. The person is wearing blue jeans. The first photograph shows the person wearing red sneakers. The second photograph shows the person wearing red sneakers. The third photograph shows the person wearing black slippers. The background is a white metal fence and a concrete floor.

Slide 3 Target l-word: 扑拉 ‘push slightly to remove the dirt or make soft things more smooth’

下面两个动图有什么区别？

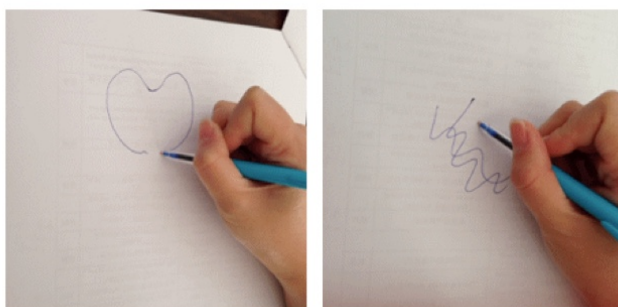
What is the difference between the two pictures?



Slide 4 Target l-word: 划拉 ‘scribble’

下面两个动图有什么区别？

What is the difference between the two pictures?



Slide 5 Target l-word: 提溜 ‘carry sth. randomly’

下面两个动图有什么区别？

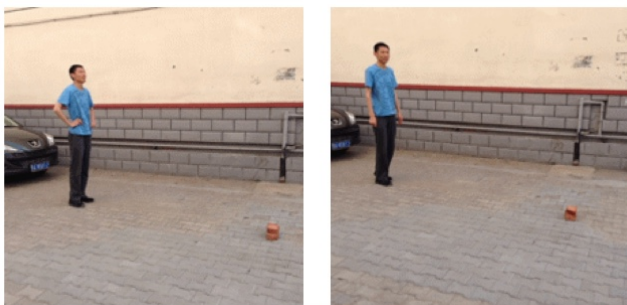
What is the difference between the two pictures?



Slide 6 Target l-word: 卜烂 ‘trip up, stumble’

下面两个动图有什么区别？

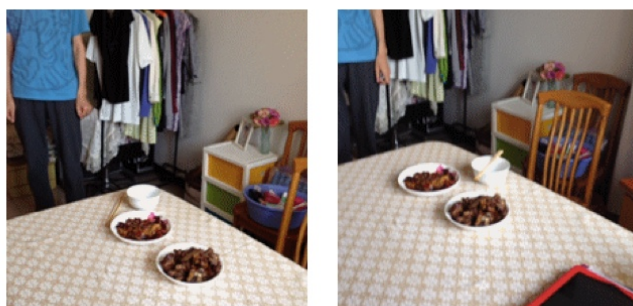
What is the difference between the two pictures?



Slide 7 Target l-word: 拨拉 ‘repeatedly move horizontally with hand or stick’

下面两个动图有什么区别？

What is the difference between the two pictures?



Slide 8 Target l-word: 圪料 ‘bended, curved’

下面两幅图有什么区别？

What is the difference between the two pictures?



Slide 9 Target l-word: 耷拉 ‘droop, hanging’

下面两幅图有什么区别？

What is the difference between the two pictures?



Slide 10 Target l-word: 嘟噜 ‘a cluster/bunch of sth.’

下面两幅图有什么区别？

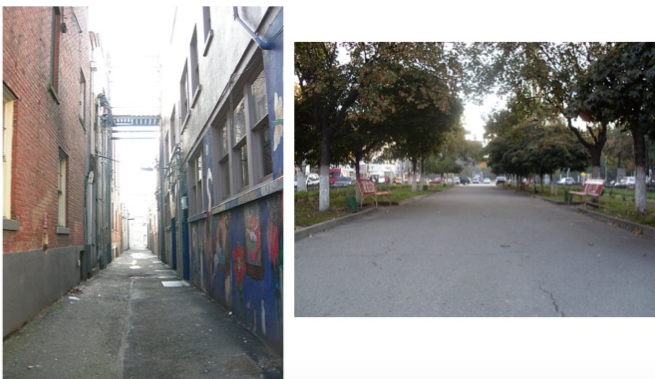
What is the difference between the two pictures?



Slide 11 Target l-word: 合浪 ‘an alley or a lane’

下面两幅图有什么区别？

What is the difference between the two pictures?



Slide 12 Target l-word: 拨拉 ‘repeatedly move horizontally with hand or stick’

下图中发生了什么？
What is happening in the picture?



Slide 13 Target l-word: 葫芦 ‘a calabash’

请描述一下右图。
Please describe
the picture.



Slide 14 Target l-word: 轱辘 ‘wheels’

下面三幅图有什么区别？
What is the difference between the three pictures?



Slide 15 Target l-word: 耷拉 'droop, hanging'

下面两幅图有什么区别？
What is the difference between the two pictures?



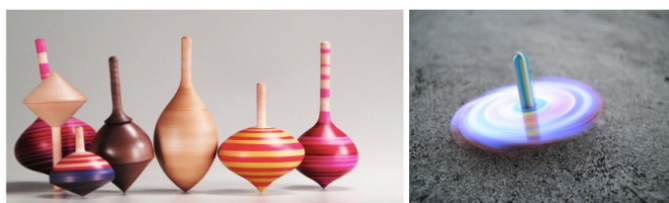
Slide 16 Target l-word: 卜浪 'an onomatopoeia for the sound of a rattle-drum'

下面两幅图有什么区别？
What is the difference between the two pictures?



Slide 17 Target l-word: 陀螺 'a spinning top'

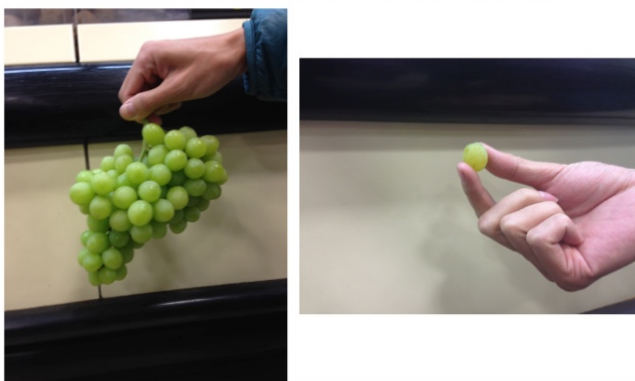
下面两幅图有什么区别？
What is the difference between the two pictures?



Slide 18 Target l-word: 嘟噜 ‘a cluster/bunch of sth.’

下面两幅图有什么区别？

What is the difference between the two pictures?



Slide 19 Target l-word: 嘟噜 ‘a cluster/bunch of sth.’

下面两幅图有什么区别？

What is the difference between the two pictures?



Slide 20 Target l-word: 窟窿 ‘holes’

下图中装苹果的塑料袋有什么特点？

What can you see on the plastic bag?



References

- carrotmadman6. (2010). *Spinning top*[Photograph]. Retrieved June 26th, 2014, from: [http://commons.wikimedia.org/wiki/File:Spinning_top_\(5448672388\).jpg](http://commons.wikimedia.org/wiki/File:Spinning_top_(5448672388).jpg)
 - Cat pushes a glass[GIF]. (2014). Retrieved June 26th, 2014, from <http://gifak-net.tumblr.com/archive>.
 - Cliff Reppart. (2011). *English Lop in the snow*[Photograph]. Retrieved June 26th, 2014, from: http://commons.wikimedia.org/wiki/File:English_Lop_Rabbit.jpg
 - David Earle. (2009). *A tough gang of Spinning Tops*[Photograph]. Retrieved June 26th, 2014, from http://en.wikipedia.org/wiki/Top#mediaviewer/File:A_tough_gang_of_Spinning_Tops.jpg.
 - Ebiebi2. (2008). *Den-den daiko, Image of a "wave drum"*[Photograph]. Retrieved June 26th, 2014, from: http://en.wikipedia.org/wiki/Pellet_drum#mediaviewer/File:Dendendaiko.jpg.
 - Giku (2008). *The Renasterii Avenue in Chisinau*[Photograph]. Retrieved June 26th, 2014, from Wikimedia commons website: http://commons.wikimedia.org/wiki/File:Grigore_Vieru_Avenue.jpg.
 - Gittout. (2009). *Alert jack rabbit*[Photograph]. Retrieved June 26th, 2014, from: http://commons.wikimedia.org/wiki/File:Jack_rabbit.tif
-
- Jerry "Woody" (2010). *Chinese coin*[Photograph]. Retrieved June 26th, 2014, from Wikimedia commons website: http://commons.wikimedia.org/wiki/File:CHINA,_QING_DYNASTY%5E_but_WHICH_EMPEROR_AND_WHEN%5E_a_Flickr_-_woody1778a.jpg
 - Joe Mabel (2007). *Alley. Downtown Olympia, Washington*[Photograph]. Retrieved June 26th, 2014, from Wikimedia commons website: http://commons.wikimedia.org/wiki/File:Oly_alley.jpg.
 - k3anlim (2009). *Six coins (Metal) in a red string is used to diminish the disastrous effect of Stars 2 and 5 (Earth)*[Photograph]. Retrieved June 26th, 2014, from Wikimedia commons website: <http://commons.wikimedia.org/wiki/File:Sixcoins.jpg>
 - Mielon (2014). *Bike*[Photograph]. Retrieved June 26th, 2014, from <http://commons.wikimedia.org/wiki/File:Bike-next.svg>
 - Nicai. (2009). *Pellet drum*[Photograph]. Retrieved June 26th, 2014, from: <https://www.flickr.com/photos/nicai>.
 - wmq06163116. (2010). *A string of calabash*[Photograph]. Retrieved June 26th, 2014, from <http://sucai.redocn.com/xiazai-108022.html>

Appendix C: The AAS questionnaires

1. Questionnaire for New Town participants

呼市与呼市方言 Hohhot and Dialects in Hohhot

A. 基本信息

A. Basic information

性别 gender: 男 male / 女 female

年龄 age: _____ 职业 occupation: _____

教育程度 level of education: _____

您父母的原籍 parents' hometown: 父亲 father _____ 母亲 mother _____

父母何时迁入呼市: _____

When did your parents move into this city:

您是否经常离开本地（半年以上）？ a. 是 b. 不是

Do you often go out of town (more than half a year)? a. Yes b. No

B. 语言能力

B. Language ability

请将您的普通话和此地话的听说能力在横线上标示出来。

Please indicate how well or badly you can speak Pǔtōnghuà and Jìn dialect by marking on the line.

例如 For example: 很好 Well ————|——— 很差 Badly

1. 我的普通话说得得 很好 Well ————|——— 很差 Badly

I can speak Pǔtōnghuà

2. 我听人说普通话 能听懂 Well ————|——— 听不懂 Badly

I can understand Pǔtōnghuà

3. 我的此地话说得得 很好 Well ————|——— 很差 Badly

I can speak Jìn dialect

4. 我听人说此地话 能听懂 Well ————|——— 听不懂 Badly

I can understand Jìn dialect

C. 下面列出了一些人对呼市或呼市人的想法或态度，您怎么看这些想法？

同意还是不同意他们的说法？他们的陈述是否符合您的情况？

C. Here are some statements about Hohhot or Hohhot people. What do you think of

them? Do you agree or disagree?

请将你对这些陈述同意（或不同意）的程度标示在该陈述下面的横线上，您标出的位置越靠近“同意”，表示您越同意上面的陈述，反之，标出的位置越接近“不同意”，代表您越反对上面的陈述。如果没有明确的意见或无法确定，可以划在横线的正中间。

Please indicate your level of agreement (or disagreement) with the following statements by marking on the straight line below. The closer you mark to “I agree”, the more you agree with the statement above it; and similarly, the closer you mark to “I disagree”, the more you disagree with the statement. If you don't have a clear opinion or you are not sure, you can mark at the middle point of the line.

例：奶豆腐很好吃。

For example: Cheese tastes nice.

同意 I agree ————— | ————— 不同意 I disagree

1. 此地话好听。

Jìn dialect sounds pleasant.

同意 I agree ————— 不同意 I disagree

2. 学说此地话很有意思。

Learning to speak Jìn dialect is very interesting.

同意 I agree ————— 不同意 I disagree

3. 在呼市听不懂此地话也不影响日常生活。

People can easily get by in Hohhot without knowing any Jìn dialect.

同意 I agree ————— 不同意 I disagree

4. 如果我在旧城找到一份工作，我会搬到旧城去住。

If I find a job in the Old Town, I'll move there.

同意 I agree ————— 不同意 I disagree

5. 我以自己是支边移民的后代而自豪。

I am proud of being a descendant of immigrants.

同意 I agree ————— 不同意 I disagree

6. 如果聊天时发现对方也是支边移民的后代，我会感到很亲切。

I'll feel close and intimate if I know someone is also a descendant of immigrants.

同意 I agree ————— 不同意 I disagree

7. 旧城人很土。

Old Town people are vulgar.

同意 I agree _____ 不同意 I disagree

8. 此地话能代表呼市的文化。

The local jìn dialect can represent Hohhot culture.

同意 I agree _____ 不同意 I disagree

9. 当有人问我是哪里人的时候，我会提到自己的父母是移民到呼市的。

When people ask me where I am from, I often mention that my parents are immigrants.

同意 I agree _____ 不同意 I disagree

10. 我喜欢在呼市生活。

I love living in Hohhot.

同意 I agree _____ 不同意 I disagree

11. 除了偶尔的购物、旅游，我很少去旧城。

Except for occasional shopping or touring, I seldom go to the Old Town.

同意 I agree _____ 不同意 I disagree

12. 如果我有机会去北京上海等大城市生活，我会选择去大城市。

If I have an opportunity to live in bigger cities like Beijing or Shanghai, I will choose to go.

同意 I agree _____ 不同意 I disagree

13. 此地话很土，很粗俗。

Jìn dialect is vulgar.

同意 I agree _____ 不同意 I disagree

14. 学说此地话很有用。

Learning to speak jìn dialect is very helpful.

同意 I agree _____ 不同意 I disagree

15. 我更愿意与其他的支边移民后代交朋友。

I prefer to make friends with other descendants of immigrants.

同意 I agree _____ 不同意 I disagree

16. 如果我孩子的男/女朋友是旧城人，我会反对他们交往。

If my child is seeing or dating someone from the Old Town, I would oppose it.

同意 I agree _____ 不同意 I disagree

17. 此地话很幽默。

Jìn dialect is humorous.

同意 I agree _____ 不同意 I disagree

18. 我今后会一直在呼市生活和工作。

I will live and work in Hohhot in the future.

同意 I agree _____ 不同意 I disagree

19. 如果有一天此地话彻底消失了，我会感到悲哀。

It would be sad if Jin dialect disappeared in the future.

同意 I agree _____ 不同意 I disagree

20. 电视或广播中应该有用此地话播报新闻的节目。

There should be some news reports in Jin dialect on TV and broadcasting.

同意 I agree _____ 不同意 I disagree

21. 电视或广播中应该有用此地话表演喜剧、小品等节目。

There should be some comedies in Jin dialect on TV and broadcasting.

同意 I agree _____ 不同意 I disagree

22. 我希望我的后代将来也在呼市生活和工作。

I hope my descendants can live and work in Hohhot in the future.

同意 I agree _____ 不同意 I disagree

23. 如果我有机会去另一个和呼市发展水平相近的城市生活，我会选择留在呼市。

If I have an opportunity to live in another city that is similar to Hohhot in economic developments, I will stay in Hohhot.

同意 I agree _____ 不同意 I disagree

2. Questionnaire for Old Town participants

呼市与呼市方言 Hohhot and Dialects in Hohhot

A. 基本信息

A. Basic information

性别 gender: 男 male / 女 female

年龄 age: _____ 职业 occupation: _____

教育程度 level of education: _____

您父母的原籍 parents' hometown: 父亲 father _____ 母亲 mother _____

父母何时迁入呼市: _____ (如一直是本地人则不用填)

When did your parents move into this city: (leave it blank if they are local)

您是否经常离开本地 (半年以上)? a. 是 b. 不是

Do you often go out of town (for more than six months)? a. Yes b. No

B. 语言能力

B. Language ability

请将您的普通话和此地话的听说能力在横线上标示出来。

Please indicate how well or badly you can speak Pǔtōnghuà and Jìn dialect by marking the line.

例如 For example: 很好 Well ————|——— 很差 Badly

1. 我的普通话说得得 很好 Well ————|——— 很差 Badly
I can speak Pǔtōnghuà

2. 我听人说普通话 能听懂 Well ————|——— 听不懂 Badly
I can understand Pǔtōnghuà

3. 我的此地话说得得 很好 Well ————|——— 很差 Badly
I can speak Jìn dialect

4. 我听人说此地话 能听懂 Well ————|——— 听不懂 Badly
I can understand Jìn dialect

C. 语言选择

C. Language choice

1. 说哪种语言对您来说最容易? a. 普通话 b. 此地话
Which language do you find easiest to speak? a. Pǔtōnghuà b. Jìn dialect

2. 您和家人常用哪种语言交流？ a. 普通话 b. 此地话
Which language do you speak at home? a. Pǔtōnghuà b. Jìn dialect
3. 您在工作单位使用哪种语言？ a. 普通话 b. 此地话
Which language do you speak at work? a. Pǔtōnghuà b. Jìn dialect

**D. 下面列出了一些人对呼市或呼市人的想法或态度，您怎么看这些想法？
同意还是不同意他们的说法？他们的陈述是否符合您的情况？**

D. Here are some statements about Hohhot or Hohhot people. What do you think of them? Do you agree or disagree?

请将你对这些陈述同意（或不同意）的程度标示在该陈述下面的横线上，您标出的位置越靠近“同意”，表示您越同意上面的陈述，反之，标出的位置越接近“不同意”，代表您越反对上面的陈述。如果没有明确的意见或无法确定，可以划在横线的正中间。

Please indicate your level of agreement (or disagreement) with the following statements by marking on the straight line below. The closer you mark to “I agree”, the more you agree with the statement above it; and similarly, the closer you mark to “I disagree”, the more you disagree with the statement. If you don’t have a clear opinion or you are not sure, you can mark at the middle point of the line.

例：奶豆腐很好吃。

For example: Cheese tastes nice.

同意 I agree ————— 不同意 I disagree

1. 此地话好听。

Jìn dialect sounds pleasant.

同意 I agree ————— 不同意 I disagree

2. 在呼市听不懂此地话也不影响日常生活。

People can easily get by in Hohhot without knowing any Jìn dialect.

同意 I agree ————— 不同意 I disagree

3. 移民应该学说此地话。

Immigrants should learn to speak Jìn dialect.

同意 I agree ————— 不同意 I disagree

4. 如果我在新城找到一份工作，我会搬到新城去住。

If I find a job in the New Town, I’ll move there.

同意 I agree ————— 不同意 I disagree

5. 我很骄傲我是旧城人。

I’m proud of being an Old Towner.

同意 I agree ————— 不同意 I disagree

6. 新城的支边移民也算是本地人。

Immigrants in the New Town are also local.

同意 I agree _____ 不同意 I disagree

7. 旧城和旧城人最能代表呼市文化。

Old Town and Old Town people can represent Hohhot culture best.

同意 I agree _____ 不同意 I disagree

8. 此地话很幽默。

Jìn dialect is humorous.

同意 I agree _____ 不同意 I disagree

9. 如果我和一个会说此地话的人说话，我一定会用此地话。

If I am talking to another Jìn speaker, I will definitely speak Jìn dialect.

同意 I agree _____ 不同意 I disagree

10. 我喜欢在呼市生活。

I love living in Hohhot.

同意 I agree _____ 不同意 I disagree

11. 如果我的子女也说此地话，我会反对。

If my children speak Jìn dialect, I would oppose it.

同意 I agree _____ 不同意 I disagree

12. 如果有一天此地话彻底消失了，我会感到悲哀。

It would be sad if Jìn dialect disappeared in the future.

同意 I agree _____ 不同意 I disagree

13. 现在旧城的改造破坏了呼市的文化。

The reconstruction in the Old Town has destroyed the culture of Hohhot.

同意 I agree _____ 不同意 I disagree

14. 如果我有机会去北京上海等大城市生活，我会选择去大城市。

If I have an opportunity to live in bigger cities like Beijing or Shanghai, I will choose to go.

同意 I agree _____ 不同意 I disagree

15. 此地话很土，很粗俗。

Jìn dialect is vulgar.

同意 I agree _____ 不同意 I disagree

16. 我希望看到我的子女和孙子女也讲此地话。

I hope to see my children and grandchildren speak local Jìn dialect.

同意 I agree _____ 不同意 I disagree

17. 此地话能代表呼市文化。

The local Jìn dialect can represent Hohhot culture.

同意 I agree _____ 不同意 I disagree

18. 我今后会一直在呼市生活和工作。

I will live and work in Hohhot in the future.

同意 I agree _____ 不同意 I disagree

19. 电视或广播中应该有用此地话播报新闻的节目。

There should be some news reports in Jìn dialect on TV and broadcasting.

同意 I agree _____ 不同意 I disagree

20. 电视或广播中应该有用此地话表演喜剧、小品等节目。

There should be some comedies in Jìn dialect on TV and broadcasting.

同意 I agree _____ 不同意 I disagree

21. 我希望我的后代将来也在呼市生活和工作。

I hope my descendants can live and work in Hohhot in the future.

同意 I agree _____ 不同意 I disagree

22. 如果我有机会去另一个和呼市发展水平相近的城市生活，我会选择留在呼市。

If I have an opportunity to live in another city that is similar to Hohhot in economic developments, I will stay in Hohhot.

同意 I agree _____ 不同意 I disagree

Appendix D: Interview topics and questions

1. Communities

Attitudes towards Hohhot communities

在呼市生活是什么样的? 你喜欢在这儿生活吗? 为什么?

Tell me about what it's like to live in Hohhot. Do you like living here? Why?

你会怎么向一个不了解呼市的人介绍这个城市?

How would you describe Hohhot to someone who didn't know it?

一个最典型的呼市人应该是什么样的?

What's a typical Hohhotian?

作为一个土生土长的呼市人, 从小在呼市长大最好和最不好的方面都是什么?

What are the best and worst things about growing up and living in Hohhot?

如果有一个外地人说呼市/新城/旧城有些地方不好, 你会不会反驳? 即使有时候你知道他说的是对的。为什么?

If an outsider was complaining about Hohhot/the New Town/the Old Town, would you defend it, even if you agreed with what they said? Why?

你对旧城人/新城人是什么印象, 有什么看法?

What is your opinion of Old Town/New Town people?

正如问卷中提到的, 作为移民/移民后代, 你对自己的身份有什么看法? 会不会有时候感到尴尬?

Like we mentioned in the questionnaire, as an immigrant/descendant of immigrants, what do you think of your own identity? Have you ever felt awkward about your identity?

Changes in the community

你认为呼市这些年来有什么样的变化? 和你小时候的呼市有什么不同?

What changes have you seen in Hohhot in your lifetime? How is it different now to when you were a child?

在过去的几十年中, 你对呼市的看法和态度有没有变化? 有哪些变化?

Has your own opinion of Hohhot changed in the past few decades? What changes?

你是否注意到呼市现在有越来越多的外地人?

Are you aware of more people coming into the community from outside?

Internal division

在呼市, 哪些标签更能显示出一群人与另一群人之间的不同? 比如民族, 语言, 宗教, 住哪个社区, 上哪个学校, 等等。

In Hohhot, what factors or labels do you think best describe the difference between different groups of people? Like ethnics, language, religion, living in which community, or going to which school?

你认为在呼市旧城和新城之间有明显的界线吗? 旧城和新城有什么不同?

Do you think there is a clear division between the New Town and the Old Town? Are you aware of any differences on different sides of the border?

你的亲戚朋友中说此地话的人多还是说普通话的人多?

Among your friends and relatives, are there more Jìn speakers or more Pǔtōnghuà speakers?

你对 2000 年新划出赛罕区有什么看法?

What is your opinion of the boundary shift of Xincheng District in 2000?

2. Language

Dialect contact situation

在你成长的过程中, 哪个阶段接触此地话/普通话最多? 与哪些人的接触让你学会了更多的此地话/普通话?

As you growing up in Hohhot, when is the time you aquired most Jìn dialect features? Through contact with whom do you learned most Jìn features?

Accent/dialect label and attitude towards own dialect

你认为自己说的话是什么口音? 你喜欢自己的口音吗?

What accent would you say you had, and do you like it?

Attitude towards dialects in the community

一个典型的说呼普/此地话的人一般是什么样的?

How do you describe a typical Jìn dialect/Hū Pǔ speaker?

新城人说的呼普是普通话吗?

Do you think Hū Pǔ is standard Pǔtōnghuà? If not, why?

Ability to recognize accent (features)/delimitations

你认为在呼市从哪个地区开始, 别人就开始和你说的话不一样了?

Where, geographically, would you say people stop talking the same as you and start sounding different?

如果你在电视上或者在外地听到一个人说此地话/呼普, 你能听出来吗? 如果能的话, 你是怎么听出来的?

Can you recognize the accent of Jìn dialect/Hū Pǔ (e.g. if heard on the radio or TV?) If so, how?

Features indexing Jìn/Hū Pǔ

有没有一些特别的音或特别的用词, 你一听就觉得这个人是说此地话/呼普的?

Are there any pronunciations or ways of saying things that you would hear and recognize the speaker as a Jìn/Hū Pǔ speaker?

Older/younger differences

你认为老年人和年轻人说的此地话/呼普有没有区别(发音是否一样, 用词是否一样)?

Do you think older and younger people talk the same here (pronounce things the same and use the same words)?

Male/female differences

你认为男的和女的说的此地话/呼普有没有区别？

Do you think there's a difference between how males and females speak here?

Accommodation

什么时候你会刻意改变自己说话的口音？为什么？

Have you ever been in a situation where you've deliberately changed the way you talk? If so, why?

当你分别和说此地话、呼普或是普通话的人交流的时候，会不会改变自己说话的方式和口音？

Do you think the way you talk could change depending on whether you were talking to someone who is a Jìn, Hū Pǔ, or Pǔtōnghuà speaker?

在结束采访之前，关于呼市和呼市方言的看法你还有什么要补充的吗？

Before we finish, is there anything that you'd like to say about Hohhot and the dialects used here that you haven't had a chance to say yet?

Appendix E: Social networks activity: the wedding guest list task

1. 现在假设你要举办一个活动来庆祝你个人的一件喜事，比如婚礼。你想要邀请所有和你关系密切的人来参加并分享你的喜悦，你会邀请哪些人？把他们列在表中第一列。并在第三列中标明他们和你的关系，比如：亲戚(QQ)、同事(TS)、同学(TX)、邻居(LJ)、其他朋友(QT)。

1. Now suppose you are going to hold a big party for celebrating your personal happy occasion, like a wedding, and you would like to invite all your acquaintances to share your happy moment. Who will you invite to this party? List their names in the first column of the form, and indicate in the third column their relationship with you like relatives(QQ¹), colleagues(TS), classmates(TX), neighbors(LJ), and other friends(QT).

注意：

- a. 这是你个人举办的活动，因此你不需要考虑你爱人或长辈要邀请的人，只从你个人的圈子出发。
- b. 你只邀请和你关系密切的，你想邀请的人，不用考虑那些你不想邀请但从人际关系上说不得不邀请的人。
- c. 你不需要给出这些人的真名，可以使用他们的昵称，假名或简单的代号，也可以在返还问卷之前**将名字这一栏剪掉**。
- d. 客人的数量没有限制，您可以根据自己的情况列，如果表格不够可以另附页。

Please note that:

- a. This party is for your personal occasion, so you don't need to consider about your partner's or your parents' guests. ONLY think of your own circle.
- b. You only need to invite your close friends or relatives –people you really want to invite. You don't need to consider those you feel like they need to be invited because of certain relationship, but actually you don't want them to come².
- c. You don't need to give the names of your guests. You can use their nicknames, pseudonyms, or last names, or simply **cut off the “name” column** before you give it back to me.
- d. There's no limit to the number of your guests. You can write as many or as few as you like, and you can use additional page.

2. 在这个聚会的准备过程中，有很多需要帮忙和询问建议的事情，比如订哪家饭店、谁帮忙接送客人、需要借用一些东西等等，那么在这些人中，有哪些是你会去找他们帮忙的？在你会寻求帮助的人后面第四栏打钩。

2. During the preparation of this party, you will need some help and advices, such as suggestions of good restaurants, shuttle guests, some stuff you need to borrow, etc. Who will you turn to for help in these invited people? Please tick in the fourth column after their names.

3. 在聚会过程中也可能发生一些麻烦事，比如：需要有人帮你照看小孩或贵重物品，结束后要有人帮你打包每一桌的食物，你忘带了一个重要的东西需要有人帮你回家去取，或者发生了一些紧急情况。那么在这些人中，你会

找哪些人来帮你做这些事？在你会寻求帮助的人后面第五栏打钩。

3. During the party, there could also be some trouble. For example: someone has to look after your baby or valuable belongings, the leftover food on every table has to be packed up, someone has to go and get something important that you left at home, or some urgent issue happens. Then among those invited, who will you turn to for help if you have these issues? Please tick in the fifth column after their names.

4. 请在这些人中说此地话的人后面第六栏中打钩。

4. Now please tick people who are Jin dialect speakers in the sixth column after their names.

5. 请在这些人中和你是同年龄段的人后面第七栏中打钩。

5. Now please tick people who belong to the same generation as you in the seventh column after their names.

6. 请在最后一栏中标出他们是新城人(XC), 旧城人(JC), 还是外地人(WD)?

6. Now please indicate in the last column whether these people are New Towner(XC), Old Towner(JC) or out-of-towner(WD)?

Notes

1. QQ is the initial of the Chinese pronunciation of relatives, which is easier for participants to write. Situations are the same for TS, TX, LJ, QT, XC, and JC.

2. This will be further clarified by some examples when I give this form to the participants. For instance, they don't need to worry that if they only invite two of their colleagues, other colleagues would be angry for not being invited.

3. The Chinese word used here is not as wide in meaning as "peer" in English. It means people who are roughly at the same age with you, or who are in the same generation.

4. Two examples are put here to direct the participants to fill in the form in the right way.



1	2	3	4	5	6	7	8
名字	序号	关系	借东西 问建议	紧急情 况帮忙	说此地话	同年龄 段	新城人 旧城人
Name	No.	Relation	borrow advices	urgent issues	Jin speaker?	peer? ³	New Town or Old Town
Bob ⁴	例	TS	√		√		JC
Mary	e. g	QQ	√	√			XC
	1						
	2						
	3						
	4						
	5						
	6						
	7						
	8						
	9						
	10						
	11						
	12						
	13						
	14						
	15						
	16						
	17						
	18						
	19						
	20						
	21						
	22						
	23						
	24						
	25						
	26						
	27						
	...						
	80						

Appendix F: Informants' social information

1. Social information and attitudinal scores for the New Town informants.

Speaker	AGE	SEX	EDU	OCCU	STSP	REL	ATTJIN	HOHORE	ATTOT	IDMIG	SOCNET
NOM1	old	M	Y	white	Y	str	-1.38	-2.49	-0.49	-0.86	0.07
NOM2	old	M	N	white	Y	str	1.41	0.82	-0.57	-0.33	0.71
NOM3	old	M	Y	white	Y	str	1.42	0.34	1.16	-1.01	0.00
NOM4	old	M	Y	white	Y	str	0.11	-0.92	0.02	-0.53	0.29
NOM5	old	M	Y	white	Y	str	1.08	0.75	0.76	-0.90	0.21
NOM6	old	M	Y	white	Y	str	-0.97	-0.42	1.16	-0.48	0.00
NOM7	old	M	Y	white	Y	str	0.51	0.55	0.27	0.30	0.07
NOF1	old	F	Y	white	Y	str	-0.39	-0.91	-0.42	-0.24	0.07
NOF2	old	F	N	blue	Y	str	-1.93	0.44	2.25	-0.63	0.67
NOF3	old	F	N	white	Y	str	-2.07	0.72	2.50	-0.78	0.14
NOF4	old	F	Y	white	Y	str	0.09	0.69	-0.66	-1.14	0.00
NOF5	old	F	Y	white	Y	str	-1.34	0.41	-1.07	0.89	0.29
NOF6	old	F	Y	white	Y	str	0.77	0.47	-0.89	-0.41	0.07
NMM1	mid	M	Y	white	Y	acqu	-0.69	-0.88	0.67	0.90	0.14
NMM2	mid	M	N	blue	Y	str	-0.37	0.08	-1.04	-1.11	0.36
NMM3	mid	M	N	blue	N	acqu	-0.18	0.83	-1.28	NA	0.50
NMM4	mid	M	N	white	N	acqu	1.05	1.35	-0.95	NA	0.36
NMF1	mid	F	N	white	Y	str	-0.57	0.58	-0.26	-0.17	0.00
NMF2	mid	F	N	blue	Y	fri	-0.17	1.12	-1.26	-1.37	0.21
NMF3	mid	F	Y	white	N	acqu	-0.94	-1.23	-0.64	-0.17	0.14
NMF4	mid	F	Y	white	N	str	0.53	1.45	-0.20	NA	0.79
NMF5	mid	F	N	white	N	str	0.40	1.62	-1.08	NA	0.43
NMF6	mid	F	Y	white	Y	acqu	-1.56	0.84	-1.49	-1.10	0.21
NYM1	yng	M	Y	white	Y	fri	-0.31	-1.32	0.23	2.82	0.42
NYM2	yng	M	Y	white	N	str	1.01	0.95	1.37	NA	0.42
NYM3	yng	M	Y	white	Y	fri	1.57	-2.18	-0.89	0.06	0.00
NYM4	yng	M	Y	white	Y	fri	0.54	-0.91	-0.14	0.80	0.00
NYM5	yng	M	Y	white	N	fri	1.31	-0.95	0.23	NA	0.36
NYF1	yng	F	Y	white	N	str	-1.19	0.00	0.95	-0.07	0.00
NYF2	yng	F	Y	white	N	str	-0.34	-1.10	1.35	1.45	0.40
NYF3	yng	F	Y	white	N	fri	1.22	0.19	0.24	0.29	0.07
NYF4	yng	F	Y	white	N	str	0.70	-0.13	-0.24	0.78	0.00
NYF5	yng	F	Y	white	N	fri	0.41	-0.09	0.04	1.79	0.29
NYF6	yng	F	Y	white	Y	str	0.02	-0.71	0.87	-0.03	0.08
NYF7	yng	F	Y	white	Y	fri	0.24	0.01	-0.52	1.25	0.17

2. Social information and attitudinal scores for the Old Town informants.

Speaker	AGE	SEX	EDU	OCCU	LANG	MOVNT	REL	ATT	SOCNET
OOM1	old	M	Y	white	bi	Y	str	0.84	0.67
OOM2	old	M	Y	white	Jin	Y	str	1.30	0.75
OOM3	old	M	Y	white	bi	Y	str	-1.71	0.67
OOM4	old	M	Y	white	Jin	Y	str	-0.82	0.50
OOM5	old	M	N	white	Jin	Y	str	0.26	0.83
OOF1	old	F	Y	white	Jin	N	str	1.09	1.00
OOF2	old	F	Y	white	Jin	Y	str	1.17	0.67
OOF3	old	F	Y	white	Jin	Y	str	-1.58	0.75
OOF4	old	F	N	blue	Jin	Y	str	1.28	1.00
OMM1	mid	M	N	blue	nonJ	Y	str	NA	0.25
OMM2	mid	M	N	blue	Jin	N	acqu	-0.66	0.92
OMM3	mid	M	N	white	bi	N	str	-1.42	1.00
OMM4	mid	M	N	blue	Jin	N	str	-0.44	1.00
OMM5	mid	M	N	blue	bi	Y	str	-1.30	0.67
OMF1	mid	F	N	white	bi	N	acqu	0.55	0.42
OMF2	mid	F	N	blue	bi	Y	fri	0.78	0.50
OMF3	mid	F	N	white	bi	N	str	0.31	0.58
OMF4	mid	F	Y	white	nonJ	Y	fri	-2.13	0.25
OMF5	mid	F	N	blue	bi	Y	str	1.55	0.83
OMF6	mid	F	N	blue	bi	Y	str	0.78	0.92
OMF7	mid	F	N	blue	Jin	N	str	0.22	0.92
OYM1	yng	M	Y	white	nonJ	N	str	0.41	0.10
OYM2	yng	M	Y	white	bi	Y	str	-0.44	0.50
OYM3	yng	M	Y	white	nonJ	Y	fri	-0.11	0.17
OYM4	yng	M	Y	white	bi	N	str	0.35	0.88
OYM5	yng	M	Y	white	bi	N	str	0.85	0.70
OYF1	yng	F	Y	white	nonJ	N	str	-0.86	0.50
OYF2	yng	F	Y	white	nonJ	Y	fri	-1.13	0.13
OYF3	yng	F	Y	white	nonJ	Y	fri	-0.03	0.00
OYF4	yng	F	Y	white	nonJ	Y	str	1.00	0.50
OYF5	yng	F	N	white	bi	Y	str	-0.21	0.25
OYF6	yng	F	Y	white	nonJ	N	str	0.12	0.40

Notes: old = older generation; mid = middle-aged generation; yng = younger generation; STSP = state-sponsored migrant; REL: relation to the interviewer; str = stranger; acqu=acquaintance; fri = friend; bi = bi-dialectal speaker; Jin = Jin speaker; nonJ = non-Jin speaker; The attitudinal scores and SOCNET scores shown in both tables were rounded to two decimal places, but they were not rounded in the statistical analysis.

Appendix G: Raw data


This table presents the token number of the stress pattern variable and the fricative variable analyzed for all 67 speakers. The percentages of the weak-strong pattern and the [x] variants are also calculated for each speaker to provide a better view of the raw data.

No	Speaker	Stress pattern variation		Fricative variation	
		token no.	pct of w-s	token no.	pct of [x]
1	NOM1	26	4%	19	41%
2	NOM2	47	100%	18	47%
3	NOM3	46	7%	33	39%
4	NOM4	40	60%	22	62%
5	NOM5	61	59%	35	41%
6	NOM6	29	0%	16	38%
7	NOM7	50	42%	21	48%
8	NOF1	7	0%	2	44%
9	NOF2	42	43%	28	55%
10	NOF3	44	0%	29	12%
11	NOF4	38	8%	22	0%
12	NOF5	20	0%	4	25%
13	NOF6	30	20%	20	14%
14	NMM1	89	55%	84	5%
15	NMM2	54	57%	36	25%
16	NMM3	40	53%	20	0%
17	NMM4	41	51%	33	21%
18	NMF1	125	57%	61	33%
19	NMF2	81	58%	58	24%
20	NMF3	83	18%	36	14%
21	NMF4	62	82%	53	6%
22	NMF5	45	71%	34	13%
23	NMF6	27	30%	32	29%
24	NYM1	71	90%	66	0%
25	NYM2	34	47%	19	13%
26	NYM3	23	43%	21	29%
27	NYM4	32	31%	23	52%
28	NYM5	29	72%	24	31%
29	NYF1	20	25%	9	0%
30	NYF2	19	58%	8	68%
31	NYF3	49	27%	35	12%

32	NYF4	40	52%	29	0%
33	NYF5	30	57%	26	0%
34	NYF6	19	5%	12	43%
35	NYF7	36	72%	22	13%
36	OOM1	6	0%	3	8%
37	OOM2	58	91%	33	76%
38	OOM3	46	98%	23	27%
39	OOM4	41	93%	29	56%
40	OOM5	23	83%	20	46%
41	OOF1	24	100%	18	52%
42	OOF2	44	95%	30	21%
43	OOF3	45	100%	39	43%
44	OOF4	37	89%	29	63%
45	OMM1	27	85%	7	92%
46	OMM2	28	100%	27	41%
47	OMM3	47	100%	39	32%
48	OMM4	44	98%	29	72%
49	OMM5	48	90%	25	70%
50	OMF1	59	86%	24	51%
51	OMF2	81	80%	76	83%
52	OMF3	66	65%	51	0%
53	OMF4	24	8%	34	45%
54	OMF5	40	93%	37	22%
55	OMF6	35	94%	33	28%
56	OMF7	51	96%	39	75%
57	OYM1	57	70%	45	11%
58	OYM2	63	67%	36	25%
59	OYM3	55	78%	54	86%
60	OYM4	19	74%	15	8%
61	OYM5	39	87%	31	52%
62	OYF1	24	63%	18	60%
63	OYF2	36	69%	12	49%
64	OYF3	36	81%	28	36%
65	OYF4	15	40%	12	35%
66	OYF5	28	64%	21	33%
67	OYF6	37	81%	30	23%
Total token no.		2812		1957	

Appendix H: Human Ethics Application

Approval for the research involving human participants described in this thesis was granted by the Human Ethics Committee of the University of Canterbury under the reference number HEC 2014/12/LR-PS. The application is presented below.

<p style="text-align: center;">UNIVERSITY OF CANTERBURY HUMAN ETHICS COMMITTEE PhD & STAFF LOW RISK APPLICATION (For research proposals which are not considered in full by the University Human Ethics Committee)</p>	
---	---

ETHICAL APPROVAL OF LOW RISK RESEARCH INVOLVING HUMAN PARTICIPANTS REVIEWED BY DEPARTMENTS

Please read the important notes appended to this form before completing the sections below

- 1 RESEARCHER'S NAME: Xuan Wang
- 2 NAME OF DEPARTMENT OR SCHOOL: Linguistics, School of Language Social and Political Sciences
- 3 EMAIL ADDRESS: xuan.wang@pg.canterbury.ac.nz
- 4 TITLE OF PROJECT: A sociophonetic account of l-words in Chinese Jin dialect
- 5 PROJECTED START DATE OF PROJECT: Aug 3rd, 2014
- 6 STAFF MEMBER/SUPERVISOR RESPONSIBLE FOR PROJECT: Dr. Kevin Watson
- 7 NAMES OF OTHER PARTICIPATING STAFF AND STUDENTS: Dr. Lynn Clark
- 8 STATUS OF RESEARCH: (pilot study, thesis, staff research – please include status of student researchers involved if this is a staff-led project) PhD thesis
- 9 BRIEF DESCRIPTION OF THE PROJECT:
Please give a brief summary (approx. 500 words) of the nature of the proposal in lay language, including the aims/objectives/hypotheses of the project, rationale, participant description, and procedures/methods of the project:

Objectives

This project is a sociolinguistic study. It examines the role of identity during the formation of a new dialect in Hohhot, an immigrant city in northern China. In Hohhot, local residents are Jin dialect speakers, while immigrants came in the 1950s and 1960s from all across the country, which led to the formation of a dialect mixture: Hohhot Mandarin. This project will explore how the local people and immigrants feel about their own identity, to what degree they integrate into the local community, and whether these factors influence how they speak. The linguistic feature examined is “l-words”, so information like whether people use l-words or not, how they pronounce l-words, will be gathered during fieldwork.

Participant description

The participants of this study will be people resident in three different districts in Hohhot, who must be either born and raised in Hohhot, or moved in to Hohhot by the government policy in 1950s and 1960s.

As Hohhot is the researcher's hometown, all the participants will be recruited by her own contacts or introduced by friends and families. The closest friends will not be included, so that all the participants would volunteer to participate. No inducement will be offered to recruit participants.

There will be equal numbers of participants who are male or female, who belong to old/middle/young generation, and who are resident in Yuquan/Xincheng/Saihan District (see Table 1). So the total number of participants would be 72.

Table 1:
Participants distribution

Residence	Yuquan District						Xincheng District						Saihan District					
Gender	M			F			M			F			M			F		
Age	O	M	Y	O	M	Y	O	M	Y	O	M	Y	O	M	Y	O	M	Y
Number of participants	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
Total	24 x 3 = 72																	

M: male F: female O: old generation M: middle generation Y: young generation

Procedures

There will be four tasks for each participant. The whole process will be recorded on a digital audio recorder after asking for permission from the participants.

1) L-words elicitation activity and wordlist reading

At the beginning, participants will be asked to describe some pictures or GIF pictures which could elicit l-words. This task is necessary because l-words do not appear frequently in natural speech, and many are verbs or with abstract meaning. Then, participants will be asked to read a wordlist.

This part will take about 20 minutes.

2) Questionnaire

The second task is a questionnaire to test participants' viewpoints about their own identity, their attitudes towards Hohhot and different dialects.

This part will take 5-10 minutes.

3) Interview

After this, the researcher will continue the topics raised in the questionnaire and begin a structured interview.

This part will take about 30 minutes.

4) Social networks activity

The final task is to make an imaginary party guest list. To protect their personal information, the researcher will ask them to cut off the guests name column at last. This task is to examine the participants' social networks, like how many of their close friends are Jin dialect speakers, or to what degree they are integrated into the local community.

This task will be given to the participants at the end of the interview, and they are asked to finish and return it to the researcher after a week.

10 WHY IS THIS A LOW RISK APPLICATION?

Description should include issues raised in the Low Risk Checklist

Please give details of any ethical issues which were identified during the consideration of the proposal and the way in which these issues were dealt with or resolved.

The study will not involve physical procedures or potential physical harm of any kind. It will not induce mental stress or anxiety beyond the risks encountered in the normal life of the participants. No drugs, placebos or other non-food substances will be administered as part of this study.

Most of the participants will be the acquaintances of the researchers' friends or family. Interviews with young group participants will take place in a quiet classroom or office at the local university, and some of them will be conducted with two participants at the same time. Interviews with middle-aged group and old group participants will probably be conducted in their office or private homes. Whenever the researcher needs to interview people in their private homes, her husband will be with her.

PhD & Staff Low Risk Application Form

The study is to investigate people's viewpoints and attitudes towards their identity and language. No sensitive or private issues will be talked about during the interview. The social networks task will be conducted anonymously and imaginarily– the name column of their guest list will be cut off before they return the list to the researcher to protect their personal information.

Since the interview questions will focus on identity, participants might be mentioning some personal issues like their childhood memories, good/bad experiences with people in different communities, experiences of different generations in their families, their opinions of hometown, etc. The researcher will respect people's willingness to respond, and not to exert too much pressure or make people feel uneasy. If any participant gets upset during the interview, I'll stop and give them a chance to decide whether or not they want to withdraw, re-schedule, or carry on. The researcher will also explicitly explain the confidentiality procedures taken in the project, and try to conduct interviews based on mutual respect, trust and rapport.

This is not cross cultural research. All the participants and the researcher herself are Hohhoters. No other cultural issues would be involved.

There is no risk that the highly sensitive nature of the research topic might lead to disclosures from the participant concerning their own involvement in illegal activities or other activities that represent a risk to themselves or others.

The study does not involve participants who are particularly vulnerable or unable to give informed consent or in a dependent position (e.g. vulnerable children, people with learning difficulties or people in care facilities). The youngest group of participants will be people born in the 1980s, so no children will be involved.

There's no conflict of interest between the researcher and the participants. The researcher is currently a PhD student at UC. She is not a lecturer, teacher, treatment-provider, or employer of any participants.

Participants will not be asked to take part in the study without their consent or knowledge at the time and no deception of any sort will be used. Recording on digital audio tape recorder will be done only after asking for permission from the participants. No inducement will be offered to any participants.

There are no ethical approval procedures needed in China for this research.

11 PROVIDE COPIES OF INFORMATION & CONSENT FORMS FOR PARTICIPANTS

These forms should be on University of Canterbury departmental letterhead. The name of the project, name(s) of researcher(s), contact details of researchers (and for PhD students, the supervisor), names of who has access to the data, the length of time the data is to be stored, that participants have the right to withdraw participation and data provided, and what the data will be used for should all be clearly stated. A statement that the project has been reviewed approved by the appropriate department and the UCHEC Low Risk Approval process should also be included.

(Information sheet and Consent form are attached below with both English and Chinese versions)

Linguistics Department
 School of Language, Social and Political Sciences
 Telephone: +64 3 364 2987 ext. 8864
 Email: xuan.wang@pg.canterbury.ac.nz
 June 30th, 2014



Hohhot and Dialects in Hohhot Information Sheet for Participants

You are asked to participate in a research study conducted by Xuan Wang as part of her PhD research at the University of Canterbury. This project is being conducted under the supervision of Dr. Kevin Watson and Dr. Lynn Clark. We are interested in how recent social changes in Hohhot are affecting the dialects spoken in the city. In this project, we will be carrying out interviews with people living in Hohhot to get a better picture of life and dialects here.

You involvement in this project will be as follows:

1. we would like to ask you to describe some pictures and read some words containing the local dialect words.
2. we would like to ask you to complete a questionnaire and ask for your opinions about life in Hohhot and your experiences of growing up/living here.

This will all be recorded on a digital tape recorder and last no longer than 90 minutes.

As a follow-up to this investigation, you will be asked to complete a guest list for your imaginary party, and return this list to me after a week.

There are no potential risks or discomforts in this study.

If you would like to receive a copy of the project results when it is complete, please tick the box at the end of the consent form and provide your e-mail address.

Participation is voluntary and you have the right to withdraw at any stage without penalty. If you withdraw, I will remove information relating to you. If you wish to have your data removed you must indicate this within 7 days after you complete the activities.

The results of the project may be published, but you may be assured of the complete confidentiality of data gathered in this investigation: your identity will not be made public without your prior consent. No publication results will allow identification of the participants and their positive/negative sentiments. To ensure anonymity and confidentiality, you will be identified by number instead of name, and all data with identifying information will be kept locked up in the Department of Linguistics at the University of Canterbury. All other data will be stored on a password protected server at the University of Canterbury. Such data will not contain any identifying information. The data will be retained for 10 years. A thesis is a public document and will be available through the UC Library.

If you have any questions or concerns about the research, please feel free to contact Xuan Wang at xuan.wang@pg.canterbury.ac.nz or Dr. Kevin Watson at kevin.watson@canterbury.ac.nz. They would be pleased to discuss any concerns you may have about participation in the project.

This project has been reviewed and approved by the University of Canterbury Human Ethics Committee, and participants should address any complaints to The Chair, Human Ethics Committee, University of Canterbury, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz).

If you agree to participate in the study, you are asked to complete the consent form and return it to me.

Xuan Wang

PhD & Staff Low Risk Application Form

Linguistics Department
 School of Language, Social and Political Sciences
 Telephone: +64 3 364 2987 ext. 8864
 Email: xuan.wang@pg.canterbury.ac.nz



Hohhot and Dialects in Hohhot Consent form for Participants

I have been given a full explanation of this project and have had the opportunity to ask questions.

I understand what is required of me if I agree to take part in the research.

I understand that participation is voluntary and I may withdraw at any time without penalty. Withdrawal of participation will also include the withdrawal of any information I have provided should this remain practically achievable. And If I wish to have my data removed, I need to indicate this within 7 days after completing the activities.

I understand that any information or opinions I provide will be kept confidential to the researcher and that any published or reported results will not identify the participants and their positive/negative sentiment. I understand that a thesis is a public document and will be available through the UC Library.

I understand that all data collected for the study will be kept in locked and secure facilities and in password protected electronic form and will be destroyed after ten years.

I understand that I can contact the researcher Xuan Wang at xuan.wang@pg.canterbury.ac.nz or supervisor Dr. Kevin Watson at kevin.watson@canterbury.ac.nz for further information. If I have any complaints, I can contact the Chair of the University of Canterbury Human Ethics Committee, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz)

By signing below, I agree to participate in this research project.

Name: _____

Date: _____

Signature: _____

☐ I would like to receive a copy of the project results. My e-mail address is _____

Xuan Wang

Linguistics Department
 School of Language, Social and Political Sciences
 Telephone: +64 3 364 2987 ext. 8864
 Email: xuan.wang@pg.canterbury.ac.nz
 June 30th, 2014



呼市与呼市方言 研究参与说明

这项研究是新西兰坎特伯雷大学语言学系博士生王璇的博士论文的一部分，指导教师为 Kevin Watson 博士和 Lynn Clark 博士。我们想要了解在呼市发生的社会变化对呼市地区的方言有什么样的影响，因此我们会对呼市的居民进行采访和调研，希望能更好的了解这里的生活和语言。

如果您愿意参加我们的研究，我们会请您做以下几件事：

1. 会请您描述一些图片或动图，然后读一个方言的词汇表；
2. 会请您做一个问卷，然后就问卷中的话题对您进行采访，主要是聊聊您对呼市的看法和在呼市长大或生活的经历。

整个过程中会进行录音，并不会超过 90 分钟。

作为后续研究的需要，我们会请您完成一个邀请朋友的游戏，假设您要举办一个个人的宴会，会怎样拟客人的名单。这个游戏会在采访结束后留给您回去独自完成，一周后我会和您联系取回名单。

参与这项研究不会给您带来潜在的风险和不适。

如果您想知道这项研究的结果，可以在同意书最后的方框里打钩并留下您的电子邮箱地址，当研究结束时我会将结果发给您。

参与这项研究是完全自愿的，您可以在参与过程中的任何阶段选择退出，那么我将会把您的信息从我的数据库中移除并永不使用。如果您决定退出，请在参与完所有活动后一周之内告知我。

这项研究的结果有可能会发表，但是所有采集的信息都一定是保密的：我们绝不会不经您的同意就将您的信息公开发表。任何研究成果的发表都不会将您的身份和您表达的积极或消极的情绪想法相关联或泄露。为了保证匿名和保密，您的名字会用数字来代替，所有含有身份信息的数据都将被封存在坎特伯雷大学语言学系。其他不含身份信息的数据会被输入坎特伯雷大学的服务器，且这个服务器是有密码保护的。所有数据会在十年后销毁。我的论文是公开的文件，可以在坎特伯雷大学图书馆找到。

如果您对这项研究有任何疑问，请联系王璇（电子邮件：xuan.wang@pg.canterbury.ac.nz）或 Kevin Watson 博士（电子邮件：xuan.wang@pg.canterbury.ac.nz）。他们会为您做详细的解答。

这项研究已经通过了坎特伯雷大学人权道德委员会的审核和批准，如有问题您可以直接向该委员会主席投诉，地址及联系方式如下：

新西兰基督城坎特伯雷大学人权道德委员会 主席收 邮编：4800

The Chair, Human Ethics Committee, University of Canterbury, Private Bag 4800, Christchurch

电子邮件：human-ethics@canterbury.ac.nz。

PhD & Staff Low Risk Application Form

如果您同意参与这项研究，请在同意书上签字，非常感谢！

王璇

2014年6月30日

个人联系方式：

新西兰坎特伯雷大学语言学系

电话：+64 3 364 2987 转 8864

电子邮件：xuan.wang@pg.canterbury.ac.nz

Linguistics Department
 School of Language, Social and Political Sciences
 Telephone: +64 3 364 2987 ext. 8864
 Email: xuan.wang@pg.canterbury.ac.nz



呼市与呼市方言 参与研究同意书

研究者已经详细的为我解释了这项研究，也给了我提问的机会。

我已经了解如果我同意参加这项研究的话，我需要做什么。

我知道参与这项研究是自愿的，而且我可以在任何时候退出，我退出后我的所有已提供的信息将被移除不再使用。如果我决定退出，需要在参与完所有活动后一周之内告知研究者。

我了解我提供的所有信息和看法都是保密的，任何出版物或文章都不可以提到我的身份或将我的身份与我表达的积极或消极的情绪想法想关联。我知道论文是公开文件，在坎特伯雷大学的图书馆可以找到。

我了解这项研究中所有的数据将被封存上锁并且输入有密码保护的服务器，十年后将被销毁。

我了解如果我对参与这项研究有任何的疑问，我可以联系研究者王璇（电子邮件 xuan.wang@pg.canterbury.ac.nz），或者她的导师 Kevin Watson 博士（电子邮件 kevin.watson@canterbury.ac.nz）。如果投诉，可以联系坎特伯雷大学人权道德委员会主席，地址是：新西兰基督城坎特伯雷大学人权道德委员会 4800。电子邮件：human-ethics@canterbury.ac.nz

我同意参与这项研究并在下面签字。

姓名: _____

日期: _____

签名: _____

☐ 我希望收到研究结果报告，我的电子邮箱地址是: _____

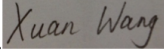
王璇

新西兰坎特伯雷大学语言学系

电话: +64 3 364 2987 转 8864

电子邮件: xuan.wang@pg.canterbury.ac.nz

Please ensure that Section A (where appropriate), B and C below are all completed

Applicant's Signature:  Date June 30th, 2014.....


A SUPERVISOR'S DECLARATION FOR PhD RESEARCH:

- 1 I have made the applicant fully aware of the need for and the requirement of seeking HEC approval for research involving human participants.
- 2 I have ensured the applicant is conversant with the procedures involved in making such an application.
- 3 In addition to this form the applicant has individually filled in the full application form which has been reviewed by me.

Signed (Supervisor): Date

B SUPPORTED BY THE DEPARTMENTAL/SCHOOL RESEARCH COMMITTEE:

NameDr Lynn Clark.....


Signature: Date..... 1.07.14.....

C APPROVED BY HEAD OF DEPARTMENT/SCHOOL:

Name

Signature: Date

SUBMISSION OF APPLICATION:

- Please attach copies of any Information Sheet and Consent Form
- Forward two hard copies to: The Secretary, Human Ethics Committee, Okeover House
- Forward an electronic copy to: human-ethics@canterbury.ac.nz

NOTES ON PROCEDURE:

The Chair of the University of Canterbury Human Ethics Committee and two other Human Ethics Committee members will review this application.

In normal circumstances queries will be forwarded via email to the applicant within 7 days

If you are a PhD student, please include a copy of this form as an appendix in your thesis

ACTION TAKEN BY HUMAN ETHICS COMMITTEE:

- ☐ Added to PhD & Staff Low Risk Reporting Database
- ☐ Referred to University of Canterbury HEC
- ☐ Referred to another Ethics Committee – please specify:

.....

REVIEWED BY: (HEC Chair)
 (HEC Member)
 (HEC Member)

Date

NOTES CONCERNING LOW RISK REPORTING SHEETS

1. This form should only be used for proposals which are Low Risk as defined in the University of Canterbury Human Ethics Committee Principles and Guidelines policy document and which may therefore be properly considered and approved at departmental level and by the Chair and two members of the University of Canterbury Human Ethics Committee under Section 5 of that document.
2. Low Risk applications are:

PhD thesis, pilot studies and staff research where the projects do not raise any issue of deception, threat, invasion of privacy, mental, physical or cultural risk or stress, and do not involve gathering personal information of a sensitive nature about or from individuals.
3. No research can be counted as low risk if it involves:
 - (i) invasive physical procedures or potential for physical harm
 - (ii) procedures which might cause mental/emotional stress or distress, moral or cultural offence
 - (iii) personal or sensitive issues
 - (iv) vulnerable groups
 - (v) Tangata Whenua
 - (vi) cross cultural research
 - (vii) investigation of illegal behaviour(s)
 - (viii) invasion of privacy
 - (ix) collection of information that might be disadvantageous to the participant
 - (x) use of information already collected that is not in the public arena which might be disadvantageous to the participant
 - (xi) use of information already collected which was collected under agreement of confidentiality
 - (xii) participants who are unable to give informed consent
 - (xiii) conflict of interest e.g. the researcher is also the lecturer, teacher, treatment-provider, colleague or employer of the research participants, or there is any other power relationship between the researcher and the research participants.
 - (xiv) deception
 - (xv) audio or visual recording without consent
 - (xvi) withholding benefits from "control" groups
 - (xvii) inducements
 - (xviii) risks to the researcher

This list is not definitive but is intended to sensitise the researcher to the types of issues to be considered. Low risk research would involve the same risk as might be encountered in normal daily life.
4. Responsibility

Supervisors are responsible for:
Theses where the projects do not raise any issues listed below.

Heads of Department are responsible for:

 - (i) Giving final approval for the low risk application.
 - (ii) Ensuring a copy of all applications are kept on file in the Department/School.

NOTE: If the HOD is the applicant, then a senior member of staff and preferably also the department and/or school research committee should give final approval. The HOD is still responsible for (ii) above.
4. A separate low risk form should be completed for each research proposal involving human participants and for which ethical approval has been considered or given at Departmental level.
5. Two completed and signed Application forms, together with a copies of Information Sheets and/ or Consent Forms, should be submitted to the Secretary, Human Ethics Committee, Okeover House, as soon as the proposal has been considered at departmental level. Please also submit an electronic version to human-ethics@canterbury.ac.nz.

PhD & Staff Low Risk Application Form

6. The Information Sheet and Consent Form include the statement "This proposal has been reviewed and approved by the University of Canterbury Human Ethics Committee low risk process".
7. Please ensure the Consent Form and the Information Sheet are on University of Canterbury letterhead and have been carefully proof-read; the institution as a whole is likely to be judged by them.
9. The research must be consistent with the University of Canterbury Human Ethics Committee Principles and Guidelines. Refer to the appendices of the University of Canterbury Human Ethics Committee Principles and Guidelines for guidance on information sheets and consent forms.
10. Please note that if the nature, procedures, location or personnel of the research project changes after departmental approval has been given in such a way that the research no longer meets the conditions laid out in Section 5 of the Principles and Guidelines, a full application to the Human Ethics Committee must be submitted.
11. This form is available electronically at: <http://www.canterbury.ac.nz/humanethics>

CHECKLIST

Please check that your application/summary has discussed:

- Procedures for voluntary, informed consent
- Privacy & confidentiality
- Risk to participants
- Obligations under the Treaty of Waitangi
- Needs of dependent persons
- Conflict of interest
- Permission for access to participants from other individuals or bodies
- Inducements

In some circumstances research which appears to meet low risk criteria may need to be reviewed by the University of Canterbury Human Ethics Committee. This might be because of requirements of:

- The publisher of the research
- An organisation which is providing funding resources, existing data, access to participants etc.
- Research which meets the criteria for review by a Health and Disability Ethics Committee – see HRC web site.

If you require advice on the appropriateness of research for low risk review, please contact the Chair of the University of Canterbury Human Ethics Committee.

Appendix I: Permission from publishers

Material covered in chapter 5 has been accepted to be published as:

Wang, X. (to appear). Investigating the role of speaker attitudes in koinéisation in Hohhot, China. *Asia Pacific Language Variation*.

See permission from the publisher below.

Dear Xuan Wang,

You have our permission to use your manuscript, ‘Investigating the role of speaker attitudes in koinéisation in Hohhot’, China’ accepted for publication in APLV, for the purposes of your Ph.D Dissertation with due acknowledgements.

With Best Wishes

Shobha Satyanath

Editor, Asia Pacific Language Variation